PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12N 15/00	A1	(11) International Publication Number: **WO 93/12227** (43) International Publication Date: 24 June 1993 (24.06.93)

(21) International Application Number: PCT/US92/10983

(22) International Filing Date: 17 December 1992 (17.12.92)

(30) Priority data:
07/810,279	17 December 1991 (17.12.91)	US
07/853,408	18 March 1992 (18.03.92)	US
07/904,068	23 June 1992 (23.06.92)	US

(60) Parent Application or Grant
(63) Related by Continuation
US	07/904,068 (CIP)
Filed on	23 June 1992 (23.06.92)

(71) Applicant (for all designated States except US): GEN-PHARM INTERNATIONAL, INC. [US/US]; 2375 Garcia Avenue, Mountain View, CA 94043 (US).

(72) Inventors; and
(75) Inventors/Applicants (for US only) : LONBERG, Nils [US/US]; 156 Carmel, San Francisco, CA 94117 (US). KAY, Robert, M. [US/US]; 2127 Broadway #5, San Francisco, CA 94115 (US).

(74) Agents: SMITH, William, M. et al.; Townsend and Townsend, One Market Plaza, 2000 Steuart Tower, San Francisco, CA 94105 (US).

(81) Designated States: AT, AU, BB, BG, BR, CA, CH, CS, DE, DK, ES, FI, GB, HU, JP, KP, KR, LK, LU, MG, MN, MW, NL, NO, NZ, PL, RO, RU, SD, SE, UA, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, SN, TD, TG).

Published
With international search report.

(54) Title: TRANSGENIC NON-HUMAN ANIMALS CAPABLE OF PRODUCING HETEROLOGOUS ANTIBODIES

(57) Abstract

The invention relates to transgenic non-human animals capable of producing heterologous antibodies and transgenic non-human animals having inactivated endogenous immunoglobulin genes. In one aspect of the invention, endogenous immunoglobulin genes are suppresed by antisense polynucleotides and/or by antiserum directed against endogenous immunoglobulins. Heterologous antibodies are encoded by immunoglobulin genes not normally found in the genome of that species of non-human animal. In one aspect of the invention, one or more transgenes containing sequences of unrearranged heterogolous human immunoglobulin heavy chains are introduced into a non-human animal thereby forming a transgenic animal capable of functionally rearranging transgenic immunoglobulin sequences and producing a repertoire of antibodies of various isotypes encoded by human immunoglobulin genes. Such heterologous human antibodies are produced in B-cells which are thereafter immortalized, e.g., by fusing with an immortalizing cell line such as a myeloma or by manipulating such B-cells by other techniques to perpetuate a cell line capable of producing a monoclonal heterologous antibody. The invention also relates to heavy and light chain immunoglobulin transgenes for making such transgenic non-human animals as well as methods and vectors for disrupting endogenous immunoglobulin loci in the transgenic animal.

1

TECHNICAL FIELD

10   The invention relates to transgenic non-human
animals capable of producing heterologous antibodies,
transgenes used to produce such transgenic animals,
transgenes capable of functionally rearranging a heterologous
D gene in V-D-J recombination, immortalized B-cells capable of
15   producing heterologous antibodies, methods and transgenes for
producing heterologous antibodies of multiple isotypes,
methods and transgenes for inactivating or suppressing
expression of endogenous immunoglobulin loci, methods and
transgenes for producing heterologous antibodies wherein a
20   variable region sequence comprises somatic mutation as
compared to germline rearranged variable region sequences, and
transgenic nonhuman animals which produce antibodies having a
human primary sequence and which bind to human antigens.

25

BACKGROUND OF THE INVENTION

           One of the major impediments facing the development
of in vivo therapeutic and diagnostic applications for
monoclonal antibodies in humans is the intrinsic
30   immunogenicity of non-human immunoglobulins. For example, when
immunocompetent human patients are administered therapeutic
doses of rodent monoclonal antibodies, the patients produce
antibodies against the rodent immunoglobulin sequences; these
human anti-mouse antibodies (HAMA) neutralize the therapeutic
35   antibodies and can cause acute toxicity. Hence, it is
desirable to produce human immunoglobulins that are reactive
with specific human antigens that are promising therapeutic
and/or diagnostic targets. However, producing human

2

immunoglobulins that bind specifically with human antigens is problematic.

The present technology for generating monoclonal antibodies involves pre-exposing, or priming, an animal
5 (usually a rat or mouse) with antigen, harvesting B-cells from that animal, and generating a library of hybridoma clones. By screening a hybridoma population for antigen binding specificity (idiotype) and also screening for immunoglobulin class (isotype), it is possible to select hybridoma clones
10 that secrete the desired antibody.

However, when present methods for generating monoclonal antibodies are applied for the purpose of generating human antibodies that have binding specificities for human antigens, obtaining B-lymphocytes which produce
15 human immunoglobulins a serious obstacle, since humans will typically not make immune responses against self-antigens.

Hence, present methods of generating human monoclonal antibodies that are specifically reactive with human antigens are clearly insufficient. It is evident that
20 the same limitations on generating monoclonal antibodies to authentic self antigens apply where non-human species are used as the source of B-cells for making the hybridoma.

The construction of transgenic animals harboring a functional heterologous immunoglobulin transgene are a method
25 by which antibodies reactive with self antigens may be produced. However, in order to obtain expression of therapeutically useful antibodies, or hybridoma clones producing such antibodies, the transgenic animal must produce transgenic B cells that are capable of maturing through the B
30 lymphocyte development pathway. Such maturation requires the presence of surface IgM on the transgenic B cells, however isotypes other than IgM are desired for therapeutic uses. Thus, there is a need for transgenes and animals harboring such transgenes that are able to undergo functional V-D-J
35 rearrangement to generate recombinational diversity and junctional diversity. Further, such transgenes and transgenic animals preferably include cis-acting sequences that facilitate isotype switching from a first isotype that is

3

required for B cell maturation to a subsequent isotype that
has superior therapeutic utility.

A number of experiments have reported the use of
transfected cell lines to determine the specific DNA sequences
5   required for Ig gene rearrangement (reviewed by Lewis and
Gellert (1989), Cell, 59, 585-588). Such reports have
identified putative sequences and concluded that the
accessibility of these sequences to the recombinase enzymes
used for rearrangement is modulated by transcription
10 (Yancopoulos and Alt (1985), Cell, 40, 271-281). The
sequences for V(D)J joining are reportedly a highly conserved,
near-palindromic heptamer and a less well conserved AT-rich
nanomer separated by a spacer of either 12 or 23 bp (Tonegawa
(1983), Nature, 302, 575-581; Hesse, et al. (1989), Genes in
15 Dev., 3, 1053-1061). Efficient recombination reportedly
occurs only between sites containing recombination signal
sequences with different length spacer regions.

Ig gene rearrangement, though studied in tissue
culture cells, has not been extensively examined in transgenic
20 mice. Only a handful of reports have been published
describing rearrangement test constructs introduced into mice
[Buchini, et al. (1987), Nature, 326, 409-411 (unrearranged
chicken λ transgene); Goodhart, et al. (1987) , Proc. Natl.
Acad. Sci. USA, 84, 4229-4233) (unrearranged rabbit κ gene);
25 and Bruggemann, et al. (1989), Proc. Natl. Acad. Sci. USA, 86,
6709-6713 (hybrid mouse-human heavy chain)]. The results of
such experiments, however, have been variable, in some cases,
producing incomplete or minimal rearrangement of the
transgene.

30   Further, a variety of biological functions of
antibody molecules are exerted by the Fc portion of molecules,
such as the interaction with mast cells or basophils through
Fcε, and binding of complement by Fcμ or Fcγ, it further is
desirable to generate a functional diversity of antibodies of
35 a given specificity by variation of isotype.

Although transgenic animals have been generated that
incorporate transgenes encoding one or more chains of a
heterologous antibody, there have been no reports of

4

heterologous transgenes that undergo successful isotype
switching. Transgenic animals that cannot switch isotypes are
limited to producing heterologous antibodies of a single
isotype, and more specifically are limited to producing an
5   isotype that is essential for B cell maturation, such as IgM
and possibly IgD, which may be of limited therapeutic utility.
Thus, there is a need for heterologous immunoglobulin
transgenes and transgenic animals that are capable of
switching from an isotype needed for B cell development to an
10 isotype that has a desired characteristic for therapeutic use.
        Based on the foregoing, it is clear that a need
exists for methods of efficiently producing heterologous
antibodies, e.g. antibodies encoded by genetic sequences of a
first species that are produced in a second species. More
15 particularly, there is a need in the art for heterologous
immunoglobulin transgenes and transgenic animals that are
capable of undergoing functional V-D-J gene rearrangement that
incorporates all or a portion of a D gene segment which
contributes to recombinational diversity. Further, there is a
20 need in the art for transgenes and transgenic animals that can
support V-D-J recombination and isotype switching so that (1)
functional B cell development may occur, and (2)
therapeutically useful heterologous antibodies may be
produced. There is also a need for a source of B cells which
25 can be used to make hybridomas that produce monoclonal
antibodies for therapeutic or diagnostic use in the particular
species for which they are designed. A heterologous
immunoglobulin transgene capable of functional V-D-J
recombination and/or capable of isotype switching could
30 fulfill these needs.
        In accordance with the foregoing object transgenic
nonhuman animals are provided which are capable of producing a
heterologous antibody, such as a human antibody.
        Further, it is an object to provide B-cells from
35 such transgenic animals which are capable of expressing
heterologous antibodies wherein such B-cells are immortalized
to provide a source of a monoclonal antibody specific for a
particular antigen.

5

In accordance with this foregoing object, it is a further object of the invention to provide hybridoma cells that are capable of producing such heterologous monoclonal antibodies.

5   Still further, it is an object herein to provide heterologous unrearranged and rearranged immunoglobulin heavy and light chain transgenes useful for producing the aforementioned non-human transgenic animals.

Still further, it is an object herein to provide
10 methods to disrupt endogenous immunoglobulin loci in the transgenic animals.

Still further, it is an object herein to provide methods to induce heterologous antibody production in the aforementioned transgenic non-human animal.

15   A further object of the invention is to provide methods to generate an immunoglobulin variable region gene segment repertoire that is used to construct one or more transgenes of the invention.

The references discussed herein are provided solely
20 for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention.

25

SUMMARY OF THE INVENTION

Transgenic nonhuman animals are provided which are capable of producing a heterologous antibody, such as a human antibody. Such heterologous antibodies may be of various
30 isotypes, including: IgG1, IgG2, IgG3, IgG4, IgM, IgA1, IgA2, IgA_{sec}, IgD, of IgE. In order for such transgenic nonhuman animals to make an immune response, it is necessary for the transgenic B cells and pre-B cells to produce surface-bound immunoglobulin, particularly of the IgM (or possibly IgD)
35 isotype, in order to effectuate B cell development and antigen-stimulated maturation. Such expression of an IgM (or IgD) surface-bound immunoglobulin is only required during the antigen-stimulated maturation phase of B cell development, and

6

mature B cells may produce other isotypes, although only a
single switched isotype may be produced at a time.

Typically, a cell of the B-cell lineage will produce
only a single isotype at a time, although cis or trans

5   alternative RNA splicing, such as occurs naturally with the μ_S
(secreted μ) and μ_M (membrane-bound μ) forms, and the μ and δ
immunoglobulin chains, may lead to the contemporaneous
expression of multiple isotypes by a single cell. Therefore,
in order to produce heterologous antibodies of multiple

10   isotypes, specifically the therapeutically useful IgG, IgA,
and IgE isotypes, it is necessary that isotype switching
occur. Such isotype switching may be classical class-
switching or may result from one or more non-classical isotype
switching mechanisms.

15   The invention provides heterologous immunoglobulin
transgenes and transgenic nonhuman animals harboring such
transgenes, wherein the transgenic animal is capable of
producing heterologous antibodies of multiple isotypes by
undergoing isotype switching. Classical isotype switching

20   occurs by recombination events which involve at least one
switch sequence region in the transgene. Non-classical
isotype switching may occur by, for example, homologous
recombination between human σ_μ and human Σ_μ sequences (δ-
associated deletion). Alternative non-classical switching

25   mechanisms, such as intertransgene and/or interchromosomal
recombination, among others, may occur and effectuate isotype
switching. Such transgenes and transgenic nonhuman animals
produce a first immunoglobulin isotype that is necessary for
antigen-stimulated B cell maturation and can switch to encode

30   and produce one or more subsequent heterologous isotypes that
have therapeutic and/or diagnostic utility. Transgenic
nonhuman animals of the invention are thus able to produce, in
one embodiment, IgG, IgA, and/or IgE antibodies that are
encoded by human immunoglobulin genetic sequences and which

35   also bind specific human antigens with high affinity.

The invention also encompasses B-cells from such
transgenic animals that are capable of expressing heterologous
antibodies of various isotypes, wherein such B-cells are

7

immortalized to provide a source of a monoclonal antibody
specific for a particular antigen. Hybridoma cells that are
derived from such B-cells can serve as one source of such
heterologous monoclonal antibodies.

5   The invention provides heterologous unrearranged and
rearranged immunoglobulin heavy and light chain transgenes
capable of undergoing isotype switching _in vivo_ in the
aforementioned non-human transgenic animals or in explanted
lymphocytes of the B-cell lineage from such transgenic
10   animals. Such isotype switching may occur spontaneously or be
induced by treatment of the transgenic animal or explanted B-
lineage lymphocytes with agents that promote isotype
switching, such as T-cell-derived lymphokines (e.g., IL-4 and
IFN_γ).

15   Still further, the invention includes methods to
induce heterologous antibody production in the aforementioned
transgenic non-human animal, wherein such antibodies may be of
various isotypes. These methods include producing an antigen-
stimulated immune response in a transgenic nonhuman animal for
20   the generation of heterologous antibodies, particularly
heterologous antibodies of a switched isotype (i.e., IgG, IgA,
and IgE).

          This invention provides methods whereby the
transgene contains sequences that effectuate isotype
25   switching, so that the heterologous immunoglobulins produced
in the transgenic animal and monoclonal antibody clones
derived from the B-cells of said animal may be of various
isotypes.

          This invention further provides methods that
30   facilitate isotype switching of the transgene, so that
switching between particular isotypes may occur at much higher
or lower frequencies or in different temporal orders than
typically occurs in germline immunoglobulin loci. Switch
regions may be grafted from various C_H genes and ligated to
35   other C_H genes in a transgene construct; such grafted switch
sequences will typically function independently of the
associated C_H gene so that switching in the transgene
construct will typically be a function of the origin of the

8

associated switch regions. Alternatively, or in combination
with switch sequences, δ-associated deletion sequences may be
linked to various C_H genes to effect non-classical switching
by deletion of sequences between two δ-associated deletion
5 sequences. Thus, a transgene may be constructed so that a
particular C_H gene is linked to a different switch sequence
and thereby is switched to more frequently than occurs when
the naturally associated switch region is used.

This invention also provides methods to determine
10 whether isotype switching of transgene sequences has occurred
in a transgenic animal containing an immunoglobulin transgene.

The invention provides immunoglobulin transgene
constructs and methods for producing immunoglobulin transgene
constructs, some of which contain a subset of germline
15 immunoglobulin loci sequences (which may include deletions).
The invention includes a specific method for facilitated
cloning and construction of immunoglobulin transgenes,
involving a vector that employs unique XhoI and SalI
restriction sites flanked by two unique NotI sites. This
20 method exploits the complementary termini of XhoI and SalI
restrictions sites and is useful for creating large constructs
by ordered concatemerization of restriction fragments in a
vector.

The transgenes of the invention include a heavy
25 chain transgene comprising DNA encoding at least one variable
gene segment, one diversity gene segment, one joining gene
segment and one constant region gene segment. The
immunoglobulin light chain transgene comprises DNA encoding at
least one variable gene segment, one joining gene segment and
30 one constant region gene segment. The gene segments encoding
the light and heavy chain gene segments are heterologous to
the transgenic non-human animal in that they are derived from,
or correspond to, DNA encoding immunoglobulin heavy and light
chain gene segments from a species not consisting of the
35 transgenic non-human animal. In one aspect of the invention,
the transgene is constructed such that the individual gene
segments are unrearranged, i.e., not rearranged so as to
encode a functional immunoglobulin light or heavy chain. Such

9

unrearranged transgenes permit recombination of the gene
segments (functional rearrangement) and expression of the
resultant rearranged immunoglobulin heavy and/or light chains
within the transgenic non-human animal when said animal is
5   exposed to antigen.

In one aspect of the invention, heterologous heavy
and light immunoglobulin transgenes comprise relatively large
fragments of unrearranged heterologous DNA. Such fragments
typically comprise a substantial portion of the C, J (and in
10   the case of heavy chain, D) segments from a heterologous
immunoglobulin locus. In addition, such fragments also
comprise a substantial portion of the variable gene segments.
In one embodiment, such transgene constructs
comprise regulatory sequences, e.g. promoters, enhancers,
15   class switch regions, recombination signals and the like,
corresponding to sequences derived from the heterologous DNA.
Alternatively, such regulatory sequences may be incorporated
into the transgene from the same or a related species of the
non-human animal used in the invention. For example, human
20   immunoglobulin gene segments may be combined in a transgene
with a rodent immunoglobulin enhancer sequence for use in a
transgenic mouse.

In a method of the invention, a transgenic non-human
animal containing germline unrearranged light and heavy
25   immunoglobulin transgenes - that undergo VDJ joining during
D-cell differentiation - is contacted with an antigen to
induce production of a heterologous antibody in a secondary
repertoire B-cell.

Also included in the invention are vectors and
30   methods to disrupt the endogenous immunoglobulin loci in the
non-human animal to be used in the invention. Such vectors
and methods utilize a transgene, preferably positive-negative
selection vector, which is constructed such that it targets
the functional disruption of a class of gene segments encoding
35   a heavy and/or light immunoglobulin chain endogenous to the
non-human animal used in the invention. Such endogenous gene
segments include diversity, joining and constant region gene
segments. In this aspect of the invention, the

10

positive-negative selection vector is contacted with at least
one embryonic stem cell of a non-human animal after which
cells are selected wherein the positive-negative selection
vector has integrated into the genome of the non-human animal
5 by way of homologous recombination. After transplantation,
the resultant transgenic non-human animal is substantially
incapable of mounting an immunoglobulin-mediated immune
response as a result of homologous integration of the vector
into chromosomal DNA. Such immune deficient non-human animals
10 may thereafter be used for study of immune deficiencies or
used as the recipient of heterologous immunoglobulin heavy and
light chain transgenes.

The invention also provides vectors, methods, and
compositions useful for suppressing the expression of one or
15 more species of immunoglobulin chain(s), without disrupting an
endogenous immunoglobulin locus. Such methods are useful for
suppressing expression of one or more endogenous
immunoglobulin chains while permitting the expression of one
or more transgene-encoded immunoglobulin chains. Unlike
20 genetic disruption of an endogenous immunoglobulin chain
locus, suppression of immunoglobulin chain expression does not
require the time-consuming breeding that is needed to
establish transgenic animals homozygous for a disrupted
endogenous Ig locus. An additional advantage of suppression
25 as compared to engognous Ig gene disruption is that, in
certain embodiments, chain suppression is reversible within an
individual animal. For example, Ig chain suppression may be
accomplished with: (1) transgenes encoding and expressing
antisense RNA that specifically hybridizes to an endogenous Ig
30 chain gene sequence, (2) antisense oligonucleotides that
specifically hybridize to an endogenous Ig chain gene
sequence, and (3) immunoglobulins that bind specifically to an
endogenous Ig chain polypeptide.

The references discussed herein are provided solely
35 for their disclosure prior to the filing date of the present
application. Nothing herein is to be construed as an
admission that the inventors are not entitled to antedate such
disclosure by virtue of prior invention.

11

BRIEF DESCRIPTION OF THE FIGURES

Fig. 1 depicts the complementarity determining regions CDR1, CDR2 and CDR3 and framework regions FR1, FR2, FR3 and FR4 in unrearranged genomic DNA and mRNA expressed from a rearranged immunoglobulin heavy chain gene,

Fig. 2 depicts the human λ chain locus,

Fig. 3 depicts the human κ chain locus,

Fig. 4 depicts the human heavy chain locus,

Fig. 5 depicts a transgene construct containing a rearranged IgM gene ligated to a 25 kb fragment that contains human γ3 and γ1 constant regions followed by a 700 bp fragment containing the rat chain 3' enhancer sequence.

Fig. 6 is a restriction map of the human κ chain locus depicting the fragments to be used to form a light chain transgene by way of in vivo homologous recombination.

Fig. 7 depicts the construction of pGP1.

Fig. 8 depicts the construction of the polylinker contained in pGP1.

Fig. 9 depicts the fragments used to construct a human heavy chain transgene of the invention.

Fig. 10 depicts the construction of pHIG1 and pCON1.

Fig. 11 depicts the human Cγ1 fragments which are inserted into pRE3 (rat enhancer 3') to form pREG2.

Fig. 12 depicts the construction of pHIG3' and PCON.

Fig. 13 depicts the fragment containing human D region segments used in construction of the transgenes of the invention.

Fig. 14 depicts the construction of pHIG2 (D segment containing plasmid).

Fig. 15 depicts the fragments covering the human Jκ and human Cκ gene segments used in constructing a transgene of the invention.

Fig. 16 depicts the structure of pEµ.

Fig. 17 depicts the construction of pKapH.

Figs. 18A through 18D depict the construction of a positive-negative selection vector for functionally disrupting the endogenous heavy chain immunoglobulin locus of mouse.

Figs. 19A through 19C depict the construction of a positive-negative selection vector for functionally disrupting the endogenous immunoglobulin light chain loci in mouse.

Figs. 20 a through e depict the structure of a kappa light chain targeting vector.

Figs. 21 a through f depict the structure of a mouse heavy chain targeting vector.

Fig. 22 depicts the map of vector pGPe.

Fig. 23 depicts the structure of vector pJM2.

Fig. 24 depicts the structure of vector pCOR1.

Fig. 25 depicts the transgene constructs for pIGM1, pHC1 and pHC2.

Fig. 26 depicts the structure of pγe2.

Fig. 27 depicts the structure of pVGE1.

Fig. 28 depicts the assay results of human Ig expression in a pHC1 transgenic mouse.

Fig. 29 depicts the structure of pJCK1.

Fig. 30 depicts the construction of a synthetic heavy chain variable region.

Fig. 31 is a schematic representation of the heavy chain minilocus constructs $pIGM_1$, pHC1, and pHC2.

Fig. 32 is a schematic representation of the heavy chain minilocus construct pIGG1 and the κ light chain minilocus construct pKC1, pKVe1, and pKC2.

Fig. 33 depicts a scheme to reconstruct functionally rearranged light chain genes.

Fig. 34 depicts serum ELISA results

Fig. 35 depicts the results of an ELISA assay of serum from 8 transgenic mice.

Fig. 36 is a schematic representation of plasmid pBCE1.

Fig. 37 depicts the immune response of transgenic mice of the present invention against KLH-DNP, by measuring IgG and IgM levels specific for KLH-DNP (37A), KLH (37B) and BSA-DNP (37C).

Fig. 38 shows ELISA data demonstrating the presence of antibodies that bind human carcinoembryonic antigen (CEA) and comprise human μ chains; each panel shows reciprocal

13

serial dilutions from pooled serum samples obtained from mice
on the indicated day following immunization.

Fig. 39 shows ELISA data demonstrating the presence
of antibodies that bind human carcinoembryonic antigen (CEA)
5   and comprise human γ chains; each panel shows reciprocal
serial dilutions from pooled serum samples obtained from mice
on the indicated day following immunization.

Fig. 40 shows aligned variable region sequences of
23 randomly-chosen cDNAs generated from mRNA obtained from
10 lymphoid tissue of HC1 transgenic mice immunized with human
carcinoembryonic antigen (CEA) as compared to the germline
transgene sequence (top line); on each line nucleotide changes
relative to germline sequence are shown above the alteration
in deduced amino acid sequence (if any); the regions
15 corresponding to heavy chain CDR1, CDR2, and CDR3 are
indicated. Non-germline encoded nucleotides are shown in
capital letters. Germline V_H251 and J_H are shown in lower case
letters. Deduced amino acid changes are given beneath
nucleotide sequences using th conventional single-letter
20 notation.

Fig. 41 shows the data from Fig. 40 in histogram
format; deduced amino acid residue position is shown as the
ordinate (left is the amino-terminal direction, right is in
the direction towards the carboxy-terminus) and frequency of
25 sequence variation is shown as the abscissa.

Fig. 42 show the nucleotide sequence of a human DNA
fragment, designated vk65.3, containing a V_κ gene segment; the
deduced amino acid sequences of the V_κ coding regions are also
shown; splicing and recombination signal sequences
30 (heptamer/nonamer) are shown boxed.

Fig. 43 show the nucleotide sequence of a human DNA
fragment, designated vk65.5, containing a V_κ gene segment; the
deduced amino acid sequences of the V_κ coding regions are also
shown; splicing and recombination signal sequences
35 (heptamer/nonamer) are shown boxed.

Fig. 44 show the nucleotide sequence of a human DNA
fragment, designated vk65.8, containing a V_κ gene segment; the
deduced amino acid sequences of the V_κ coding regions are also

14

shown; splicing and recombination signal sequences
(heptamer/nonamer) are shown boxed.

Fig. 45 show the nucleotide sequence of a human DNA
fragment, designated vk65.15, containing a V_κ gene segment;
5 the deduced amino acid sequences of the V_κ coding regions are
also shown; splicing and recombination signal sequences
(heptamer/nonamer) are shown boxed.

Fig. 46 shows formation of a light chain minilocus
by homologous recombination between two overlapping fragments
10 which were co-injected.

Table 1 depicts the sequence of vector pGPe.

Table 2 depicts the sequence of gene $V_H4\ldots.8$.

Table 3 depicts the detection of human IgM and IgG
in the serum of transgenic mice of this invention.

15   Table 4 depicts sequences of VDJ joints.

Table 5 depicts the distribution of J segments
incorporated into pHC1 transgene encoded transcripts to J
segments found in adult human peripheral blood lymphocytes
(PBL).

20   Table 6 depicts the distribution of D segments
incorporated into pHC1 transgene encoded transcripts to D
segments found in adult human peripheral blood lymphocytes
(PBL).

Table 7 depicts the length of the CDR3 peptides from
25 transcripts with in-frame VDJ joints in the pHC1 transgenic
mouse and in human PBL.

Table 8 depicts the predicted amino acid sequences
of the VDJ regions from 30 clones analyzed from a pHC1
transgenic.

30   Table 9 shows transgenic mice of line 112 that were
used in the indicated experiments; (+) indicates the presence
of the respective transgene, (++) indicates that the animal is
homozygous for the J_HD knockout transgene.

DETAILED DESCRIPTION

35   As has been discussed supra, it is desirable to
produce human immunoglobulins that are reactive with specific
human antigens that are promising therapeutic and/or

15

diagnostic targets. However, producing human immunoglobulins
that bind specifically with human antigens is problematic.

First, the immunized animal that serves as the
source of B cells must make an immune response against the
5   presented antigen. In order for an animal to make an immune
response, the antigen presented must be foreign and the animal
must not be tolerant to the antigen. Thus, for example, if it
is desired to produce a human monoclonal antibody with an
idiotype that binds to a human protein, self-tolerance will
10 prevent an immunized human from making a substantial immune
response to the human protein, since the only epitopes of the
antigen that may be immunogenic will be those that result from
polymorphism of the protein within the human population
(allogeneic epitopes).

15   Second, if the animal that serves as the source of
B-cells for forming a hybridoma (a human in the illustrative
given example) does make an immune response against an
authentic self antigen, a severe autoimmune disease may result
in the animal. Where humans would be used as a source of B-
20 cells for a hybridoma, such autoimmunization would be
considered unethical by contemporary standards.

One methodology that can be used to obtain human
antibodies that are specifically reactive with human antigens
is the production of a transgenic mouse harboring the human
25 immunoglobulin transgene constructs of this invention.
Briefly, transgenes containing all or portions of the human
immunoglobulin heavy and light chain loci, or transgenes
containing synthetic "miniloci" (described infra, and in
PCT/US91/06185 filed August 28, 1991) which comprise essential
30 functional elements of the human heavy and light chain loci,
are employed to produce a transgenic nonhuman animal. Such a
transgenic nonhuman animal will have the capacity to produce
immunoglobulin chains that are encoded by human immunoglobulin
genes, and additionally will be capable of making an immune
35 response against human antigens. Thus, such transgenic animals
can serve as a source of immune sera reactive with specified
human antigens, and B-cells from such transgenic animals can
be fused with myeloma cells to produce hybridomas that secrete

monoclonal antibodies that are encoded by human immunoglobulin genes and which are specifically reactive with human antigens.

The production of transgenic mice containing various forms of immunoglobulin genes has been reported previously.

5   Rearranged mouse immunoglobulin heavy or light chain genes have been used to produce transgenic mice. In addition, functionally rearranged human Ig genes including the μ or $\gamma 1$ constant region have been expressed in transgenic mice. However, experiments in which the transgene comprises

10 unrearranged (V-D-J or V-J not rearranged) immunoglobulin genes have been variable, in some cases, producing incomplete or minimal rearrangement of the transgene. However, there are no published examples of either rearranged or unrearranged immunoglobulin transgenes which undergo successful isotype

15 switching between C_H genes within a transgene.

Definitions

As used herein, the term "antibody" refers to a

20 glycoprotein comprising at least two light polypeptide chains and two heavy polypeptide chains. Each of the heavy and light polypeptide chains contains a variable region (generally the amino terminal portion of the polypeptide chain) which contains a binding domain which interacts with antigen. Each

25 of the heavy and light polypeptide chains also comprises a constant region of the polypeptide chains (generally the carboxyl terminal portion) which may mediate the binding of the immunoglobulin to host tissues or factors including various cells of the immune system, some phagocytic cells and

30 the first component (C1q) of the classical complement system.

As used herein, a "heterologous antibody" is defined in relation to the transgenic non-human organism producing such an antibody. It is defined as an antibody having an amino acid sequence or an encoding DNA sequence corresponding to

35 that found in an organism not consisting of the transgenic non-human animal.

As used herein, a "heterohybrid antibody" refers to an antibody having a light and heavy chains of different

17

organismal origins. For example, an antibody having a human
heavy chain associated with a murine light chain is a
heterohybrid antibody.

5   As used herein, "isotype" refers to the antibody
class (e.g., IgM or IgG_1) that is encoded by heavy chain
constant region genes.

                As used herein, "isotype switching" refers to the
phenomenon by which the class, or isotype, of an antibody
changes from one Ig class to one of the other Ig classes.

10   As used herein, "nonswitched isotype" refers to the
isotypic class of heavy chain that is produced when no isotype
switching has taken place; the C_H gene encoding the
nonswitched isotype is typically the first C_H gene immediately
downstream from the functionally rearranged VDJ gene.

15   As used herein, the term "switch sequence" refers to
those DNA sequences responsible for switch recombination. A
"switch donor" sequence, typically a μ switch region, will be
5' (i.e., upstream) of the construct region to be deleted
during the switch recombination. The "switch acceptor" region
20 will be between the construct region to be deleted and the
replacement constant region (e.g., γ, ϵ, etc.). As there is
no specific site where recombination always occurs, the final
gene sequence will typically not be predictable from the
construct.

25   As used herein, "glycosylation pattern" is defined
as the pattern of carbohydrate units that are covalently
attached to a protein, more specifically to an immunoglobulin
protein. A glycosylation pattern of a heterologous antibody
can be characterized as being substantially similar to
30 glycosylation patterns which occur naturally on antibodies
produced by the species of the nonhuman transgenic animal,
when one of ordinary skill in the art·would recognize the
glycosylation pattern of the heterologous antibody as being
more similar to said pattern of glycosylation in the species
35 of the nonhuman transgenic animal than to the species from
which the C_H genes of the transgene were derived.

                As used herein, "specific binding" refers to the
property of the antibody: (1) to bind to a predetermined

18

antigen with an affinity of at least 1×10^7 M^{-1}, and (2) to preferentially bind to the predetermined antigen with an affinity that is at least two-fold greater than its affinity for binding to a non-specific antigen (e.g., BSA, casein) other than the predetermined antigen.

The term "naturally-occurring" as used herein as applied to an object refers to the fact that an object can be found in nature. For example, a polypeptide or polynucleotide sequence that is present in an organism (including viruses) that can be isolated from a source in nature and which has not been intentionally modified by man in the laboratory is naturally-occurring.

The term "rearranged" as used herein refers to a configuration of a heavy chain or light chain immunoglobulin locus wherein a V segment is positioned immediately adjacent to a D-J or J segment in a conformation encoding essentially a complete V_H or V_L domain, respectively. A rearranged immunoglobulin gene locus can be identified by comparison to germline DNA; a rearranged locus will have at least one recombined heptamer/nonamer homology element.

The term "unrearranged" or "germline configuration" as used herein in reference to a V segment refers to the configuration wherein the V segment is not recombined so as to be immediately adjacent to a D or J segment.

Transgenic Nonhuman Animals Capable
of Producing Heterologous Antibodies

The design of a transgenic non-human animal that responds to foreign antigen stimulation with a heterologous antibody repertoire, requires that the heterologous immunoglobulin transgenes contained within the transgenic animal function correctly throughout the pathway of B-cell development. In a preferred embodiment, correct function of a heterologous heavy chain transgene includes isotype switching. Accordingly, the transgenes of the invention are constructed so as to produce isotype switching and one or more of the following: (1) high level and cell-type specific expression, (2) functional gene rearrangement, (3) activation of and

response to allelic exclusion, (4) expression of a sufficient
primary repertoire, (5) signal transduction, (6) somatic
hypermutation, and (7) domination of the transgene antibody
locus during the immune response.

5   As will be apparent from the following disclosure,
not all of the foregoing criteria need be met. For example, in
those embodiments wherein the endogenous immunoglobulin loci
of the transgenic animal are functionally disrupted, the
transgene need not activate allelic exclusion. Further, in
10   those embodiments wherein the transgene comprises a
functionally rearranged heavy and/or light chain
immunoglobulin gene, the second criteria of functional gene
rearrangement is unnecessary, at least for that transgene
which is already rearranged. For background on molecular
15   immunology, see, Fundamental Immunology, 2nd edition (1989),
Paul William E., ed. Raven Press, N.Y., which is incorporated
herein by reference.

           In one aspect of the invention, transgenic non-human
animals are provided that contain rearranged, unrearranged or
20   a combination of rearranged and unrearranged heterologous
immunoglobulin heavy and light chain transgenes in the
germline of the transgenic animal. Each of the heavy chain
transgenes comprises at least one C_H gene. In addition, the
heavy chain transgene may contain functional isotype switch
25   sequences, which are capable of supporting isotype switching
of a heterologous transgene encoding multiple C_H genes in B-
cells of the transgenic animal. Such switch sequences may be
those which occur naturally in the germline immunoglobulin
locus from the species that serves as the source of the
30   transgene C_H genes, or such switch sequences may be derived
from those which occur in the species that is to receive the
transgene construct (the transgeneic animal). For example, a
human transgene construct that is used to produce a transgenic
mouse may produce a higher frequency of isotype switching
35   events if it incorporates switch sequences similar to those
that occur naturally in the mouse heavy chain locus, as
presumably the mouse switch sequences are optimized to
function with the mouse switch recombinase enzyme system,

whereas the human switch sequences are not. Switch sequences
made be isolated and cloned by conventional cloning methods,
or may be synthesized de novo from overlapping synthetic
oligonucleotides designed on the basis of published sequence

5   information relating to immunoglobulin switch region sequences
(Mills et al., Nucl. Acids Res. 18:7305-7316 (1991);
Sideras et al., Intl. Immunol. 1:631-642 (1989), which are
incorporated herein by reference).

        For each of the foregoing transgenic animals,
10   functionally rearranged heterologous heavy and light chain
immunoglobulin transgenes are found in a significant fraction
of the B-cells of the transgenic animal (at least 10 percent).

        The transgenes of the invention include a heavy
chain transgene comprising DNA encoding at least one variable
15   gene segment, one diversity gene segment, one joining gene
segment and at least one constant region gene segment. The
immunoglobulin light chain transgene comprises DNA encoding at
least one variable gene segment, one joining gene segment and
at least one constant region gene segment. The gene segments
20   encoding the light and heavy chain gene segments are
heterologous to the transgenic non-human animal in that they
are derived from, or correspond to, DNA encoding
immunoglobulin heavy and light chain gene segments from a
species not consisting of the transgenic non-human animal. In
25   one aspect of the invention, the transgene is constructed such
that the individual gene segments are unrearranged, i.e., not
rearranged so as to encode a functional immunoglobulin light
or heavy chain. Such unrearranged transgenes support
recombination of the V, D, and J gene segments (functional
30   rearrangement) and preferably support incorporation of all or
a portion of a D region gene segment in the resultant
rearranged immunoglobulin heavy chain within the transgenic
non-human animal when exposed to antigen.

        In an alternate embodiment, the transgenes comprise
35   an unrearranged "mini-locus". Such transgenes typically
comprise a substantial portion of the C, D, and J segments as
well as a subset of the V gene segments. In such transgene
constructs, the various regulatory sequences, e.g. promoters,

enhancers, class switch regions, splice-donor and splice-
acceptor sequences for RNA processing, recombination signals
and the like, comprise corresponding sequences derived from
the heterologous DNA. Such regulatory sequences may be
5 incorporated into the transgene from the same or a related
species of the non-human animal used in the invention. For
example, human immunoglobulin gene segments may be combined in
a transgene with a rodent immunoglobulin enhancer sequence for
use in a transgenic mouse. Alternatively, synthetic regulatory
10 sequences may be incorporated into the transgene, wherein such
synthetic regulatory sequences are not homologous to a
functional DNA sequence that is known to occur naturally in
the genomes of mammals. Synthetic regulatory sequences are
designed according to consensus rules, such as, for example,
15 those specifying the permissible sequences of a splice-
acceptor site or a promoter/enhancer motif.

The invention also includes transgenic animals
containing germ line cells having a heavy and light transgene
wherein one of the said transgenes contains rearranged gene
20 segments with the other containing unrearranged gene segments.
In the preferred embodiments, the rearranged transgene is a
light chain immunoglobulin transgene and the unrearranged
transgene is a heavy chain immunoglobulin transgene.

25 The Structure and Generation of Antibodies

The basic structure of all immunoglobulins is based
upon a unit consisting of two light polypeptide chains and two
heavy polypeptide chains. Each light chain comprises two
30 regions known as the variable light chain region and the
constant light chain region. Similarly, the immunoglobulin
heavy chain comprises two regions designated the variable
heavy chain region and the constant heavy chain region.

The constant region for the heavy or light chain is
35 encoded by genomic sequences referred to as heavy or light
constant region gene (C_H) segments. The use of a particular
heavy chain gene segment defines the class of immunoglobulin.
For example, in humans, the μ constant region gene segments

22

define the IgM class of antibody whereas the use of a γ, $\gamma2$, $\gamma3$ or $\gamma4$ constant region gene segment defines the IgG class of antibodies as well as the IgG subclasses IgG1 through IgG4. Similarly, the use of a α_1 or α_2 constant region gene segment

5   defines the IgA class of antibodies as well as the subclasses IgA1 and IgA2. The δ and ϵ constant region gene segments define the IgD and IgE antibody classes, respectively.

The variable regions of the heavy and light immunoglobulin chains together contain the antigen binding

10   domain of the antibody. Because of the need for diversity in this region of the antibody to permit binding to a wide range of antigens, the DNA encoding the initial or primary repertoire variable region comprises a number of different DNA segments derived from families of specific variable region

15   gene segments. In the case of the light chain variable region, such families comprise variable (V) gene segments and joining (J) gene segments. Thus, the initial variable region of the light chain is encoded by one V gene segment and one J gene segment each selected from the family of V and J gene

20   segments contained in the genomic DNA of the organism. In the case of the heavy chain variable region, the DNA encoding the initial or primary repertoire variable region of the heavy chain comprises one heavy chain V gene segment, one heavy chain diversity (D) gene segment and one J gene segment, each

25   selected from the appropriate V, D and J families of immunoglobulin gene segments in genomic DNA.

In order to increase the diversity of sequences that contribute to forming antibody binding sites, it is preferable that a heavy chain transgene include cis-acting sequences that

30   support functional V-D-J rearrangement that can incorporate all or part of a D region gene sequence in a rearranged V-D-J gene sequence. Typically, at least about 1 percent of expressed transgene-encoded heavy chains (or mRNAs) include recognizable D region sequences in the V region. Preferably,

35   at least about 10 percent of transgene-encoded V regions include recognizable D region sequences, more preferably at least about 30 percent, and most preferably more than 50 percent include recognizable D region sequences.

23

A recognizable D region sequence is generally at least about eight consecutive nucleotides corresponding to a sequence present in a D region gene segment of a heavy chain transgene and/or the amino acid sequence encoded by such D

5   region nucleotide sequence. For example, if a transgene includes the D region gene DHQ52, a transgene-encoded mRNA containing the sequence 5'-TAACTGGG-3' located in the V region between a V gene segment sequence and a J gene segment sequence is recognizable as containing a D region sequence,

10 specifically a DHQ52 sequence. Similarly, for example, if a transgene includes the D region gene DHQ52, a transgene-encoded heavy chain polypeptide containing the amino acid sequence -DAF- located in the V region between a V gene segment amino acid sequence and a J gene segment amino acid

15 sequence is recognizable as containing a D region sequence, specifically a DHQ52 sequence.

However, because of somatic mutation and N-region addition, some D region sequences may be recognizable but may not correspond identically to a consecutive D region sequence

20 in the transgene. For example, a nucleotide sequence 5'-CTAAXTGGGG-3', where X is A, T, or G, and which is located in a heavy chain V region and flanked by a V region gene sequence and a J region gene sequence, can be recognized as corresponding to the DHQ52 sequence 5'-CTAACTGGG-3'.

25 Similarly, for example, the polypeptide sequences -DAFDI-, -DYFDY-, or -GAFDI- located in a V region and flanked on the amino-terminal side by an amino acid sequence encoded by a transgene V gene sequence and flanked on the carboxyterminal side by an amino acid sequence encoded by a transgene J gene

30 sequence is recognizable as a D region sequence.

Therefore, because somatic mutation and N-region addition can produce mutations in sequences derived from a transgene D region, the following definition is provided as a guide for determining the presence of a recognizable D region

35 sequence. An amino acid sequence or nucleotide sequence is recognizable as a D region sequence if: (1) the sequence is located in a V region and is flanked on one side by a V gene sequence (nucleotide sequence or deduced amino acid sequence)

and on the other side by a J gene sequence (nucleotide
sequence or deduced amino acid sequence) and (2) the sequence
is substantially identical or substantially similar to a known
D gene sequence (nucleotide sequence or encoded amino acid
5 sequence).

The term "substantial identity" as used herein
denotes a characteristic of a polypeptide sequence or nucleic
acid sequence, wherein the polypeptide sequence has at least
50 percent sequence identity compared to a reference sequence,
10 and the nucleic acid sequence has at least 70 percent sequence
identity compared to a reference sequence. The percentage of
sequence identity is calculated excluding small deletions or
additions which total less than 35 percent of the reference
sequence. The reference sequence may be a subset of a larger
15 sequence, such as an entire D gene; however, the reference
sequence is at least 8 nucleotides long in the case of
polynucleotides, and at least 3 amino residues long in the
case of a polypeptide. Typically, the reference sequence is
at least 8 to 12 nucleotides or at least 3 to 4 amino acids,
20 and preferably the reference sequence is 12 to 15 nucleotides
or more, or at least 5 amino acids.

The term "substantial similarity" denotes a
characteristic of an polypeptide sequence, wherein the
polypeptide sequence has at least 80 percent similarity to a
25 reference sequence. The percentage of sequence similarity is
calculated by scoring identical amino acids or positional
conservative amino acid substitutions as similar. A
positional conservative amino acid substitution is one that
can result from a single nucleotide substitution; a first
30 amino acid is replaced by a second amino acid where a codon
for the first amino acid and a codon for the second amino acid
can differ by a single nucleotide substitution. Thus, for
example, the sequence -Lys-Glu-Arg-Val- is substantially
similar to the sequence -Asn-Asp-Ser-Val-, since the codon
35 sequence -AAA-GAA-AGA-GUU- can be mutated to -AAC-GAC-AGC-GUU-
by introducing only 3 substitution mutations, single
nucleotide substitutions in three of the four original codons.
The reference sequence may be a subset of a larger sequence,

25

such as an entire D gene; however, the reference sequence is
at least 4 amino residues long. Typically, the reference
sequence is at least 5 amino acids, and preferably the
reference sequence is 6 amino acids or more.

5

The Primary Repertoire

        The process for generating DNA encoding the heavy
and light chain immunoglobulin genes occurs primarily in
10   developing B-cells. Prior to the joining of various
immunoglobulin gene segments, the V, D, J and constant (C)
gene segments are found, for the most part, in clusters of V,
D, J and C gene segments in the precursors of primary
repertoire B-cells. Generally, all of the gene segments for a
15   heavy or light chain are located in relatively close proximity
on a single chromosome. Such genomic DNA prior to
recombination of the various immunoglobulin gene segments is
referred to herein as "unrearranged" genomic DNA. During
B-cell differentiation, one of each of the appropriate family
20   members of the V, D, J (or only V and J in the case of light
chain genes) gene segments are recombined to form functionally
rearranged heavy and light immunoglobulin genes. Such
functional rearrangement is of the variable region segments to
form DNA encoding a functional variable region. This gene
25   segment rearrangement process appears to be sequential.
First, heavy chain D-to-J joints are made, followed by heavy
chain V-to-DJ joints and light chain V-to-J joints. The DNA
encoding this initial form of a functional variable region in
a light and/or heavy chain is referred to as "functionally
30   rearranged DNA" or "rearranged DNA". In the case of the heavy
chain, such DNA is referred to as "rearranged heavy chain DNA"
and in the case of the light chain, such DNA is referred to as
"rearranged light chain DNA". Similar language is used to
describe the functional rearrangement of the transgenes of the
35   invention.
        The recombination of variable region gene segments
to form functional heavy and light chain variable regions is
mediated by recombination signal sequences (RSS's) that flank

recombinationally competent V, D and J segments. RSS's necessary and sufficient to direct recombination, comprise a dyad-symmetric heptamer, an AT-rich nonamer and an intervening spacer region of either 12 or 23 base pairs. These signals

5   are conserved among the different loci and species that carry out D-J (or V-J) recombination and are functionally interchangeable. See Oettinger, et al. (1990), Science, 248, 1517-1523 and references cited therein. The heptamer comprises the sequence CACAGTG or its analogue followed by a

10 spacer of unconserved sequence and then a nonamer having the sequence ACAAAAACC or its analogue. These sequences are found on the J, or downstream side, of each V and D gene segment. Immediately preceding the germline D and J segments are again two recombination signal sequences, first the nonamer and then

15 the heptamer again separated by an unconserved sequence. The heptameric and nonameric sequences following a V_L, V_H or D segment are complementary to those preceding the J_L, D or J_H segments with which they recombine. The spacers between the heptameric and nonameric sequences are either 12 base pairs

20 long or between 22 and 24 base pairs long.

        In addition to the rearrangement of V, D and J segments, further diversity is generated in the primary repertoire of immunoglobulin heavy and light chain by way of variable recombination between the V and J segments in the

25 light chain and between the D and J segments of the heavy chain. Such variable recombination is generated by variation in the exact place at which such segments are joined. Such variation in the light chain typically occurs within the last codon of the V gene segment and the first codon of the J

30 segment. Similar imprecision in joining occurs on the heavy chain chromosome between the D and J_H segments and may extend over as many as 10 nucleotides. Furthermore, several nucleotides may be inserted between the D and J_H and between the V_H and D gene segments which are not encoded by genomic

35 DNA. The addition of these nucleotides is known as N-region diversity.

        After VJ and/or VDJ rearrangement, transcription of the rearranged variable region and one or more constant region

gene segments located downstream from the rearranged variable
region produces a primary RNA transcript which upon
appropriate RNA splicing results in an mRNA which encodes a
full length heavy or light immunoglobulin chain. Such heavy
5   and light chains include a leader signal sequence to effect
secretion through and/or insertion of the immunoglobulin into
the transmembrane region of the B-cell. The DNA encoding this
signal sequence is contained within the first exon of the V
segment used to form the variable region of the heavy or light
10 immunoglobulin chain. Appropriate regulatory sequences are
also present in the mRNA to control translation of the mRNA to
produce the encoded heavy and light immunoglobulin
polypeptides which upon proper association with each other
form an antibody molecule.

15   The net effect of such rearrangements in the
variable region gene segments and the variable recombination
which may occur during such joining, is the production of a
primary antibody repertoire. Generally, each B-cell which has
differentiated to this stage, produces a single primary
20 repertoire antibody. During this differentiation process,
cellular events occur which suppress the functional
rearrangement of gene segments other than those contained
within the functionally rearranged Ig gene. The process by
which diploid B-cells maintain such mono-specificity is termed
25 allelic exclusion.

The Secondary Repertoire

        B-cell clones expressing immunoglobulins from within
30 the set of sequences comprising the primary repertoire are
immediately available to respond to foreign antigens. Because
of the limited diversity generated by simple VJ and VDJ
joining, the antibodies produced by the so-called primary
response are of relatively low affinity. Two different types
35 of B-cells make up this initial response: precursors of
primary antibody-forming cells and precursors of secondary
repertoire B-cells (Linton et al., Cell 59:1049-1059 (1989)).
The first type of B-cell matures into IgM-secreting plasma

cells in response to certain antigens. The other B-cells
respond to initial exposure to antigen by entering a T-cell
dependent maturation pathway.

During the T-cell dependent maturation of antigen
5   stimulated B-cell clones, the structure of the antibody
molecule on the cell surface changes in two ways: the constant
region switches to a non-IgM subtype and the sequence of the
variable region can be modified by multiple single amino acid
substitutions to produce a higher affinity antibody molecule.

10   As previously indicated, each variable region of a
heavy or light Ig chain contains an antigen binding domain.
It has been determined by amino acid and nucleic acid
sequencing that somatic mutation during the secondary response
occurs throughout the V region including the three
15   complementary determining regions (CDR1, CDR2 and CDR3) also
referred to as hypervariable regions 1, 2 and 3 (Kabat et al.
Sequences of Proteins of Immunological Interest (1991) U.S.
Department of Health and Human Services, Washington, DC,
incorporated herein by reference. The CDR1 and CDR2 are
20   located within the variable gene segment whereas the CDR3 is
largely the result of recombination between V and J gene
segments or V, D and J gene segments. Those portions of the
variable region which do not consist of CDR1, 2 or 3 are
commonly referred to as framework regions designated FR1, FR2,
25   FR3 and FR4. See Fig. 1. During hypermutation, the
rearranged DNA is mutated to give rise to new clones with
altered Ig molecules. Those clones with higher affinities for
the foreign antigen are selectively expanded by helper
T-cells, giving rise to affinity maturation of the expressed
30   antibody. Clonal selection typically results in expression of
clones containing new mutation within the CDR1, 2 and/or 3
regions. However, mutations outside these regions also occur
which influence the specificity and affinity of the antigen
binding domain.

35

40

Transgenic Non-Human Animals Capable
of Producing Heterologous Antibody

5   Transgenic non-human animals in one aspect of the
invention are produced by introducing at least one of the
immunoglobulin transgenes of the invention (discussed
hereinafter) into a zygote or early embryo of a non-human
animal. The non-human animals which are used in the invention
generally comprise any mammal which is capable of rearranging
10   immunoglobulin gene segments to produce a primary antibody
response. Such nonhuman transgenic animals may include, for
example, transgenic pigs, transgenic rats, transgenic rabbits,
transgenic cattle, and other transgenic animal species,
particularly mammalian species, known in the art. A
15   particularly preferred non-human animal is the mouse or other
members of the rodent family.

However, the invention is not limited to the use of
mice. Rather, any non-human mammal which is capable of
mounting a primary and secondary antibody response may be
20   used. Such animals include non-human primates, such as
chimpanzee, bovine, ovine, and porcine species, other members
of the rodent family, e.g. rat, as well as rabbit and guinea
pig. Particular preferred animals are mouse, rat, rabbit and
guinea pig, most preferably mouse.

25   In one embodiment of the invention, various gene
segments from the human genome are used in heavy and light
chain transgenes in an unrearranged form. In this embodiment,
such transgenes are introduced into mice. The unrearranged
gene segments of the light and/or heavy chain transgene have
30   DNA sequences unique to the human species which are
distinguishable from the endogenous immunoglobulin gene
segments in the mouse genome. They may be readily detected in
unrearranged form in the germ line and somatic cells not
consisting of B-cells and in rearranged form in B-cells.

35   In an alternate embodiment of the invention, the
transgenes comprise rearranged heavy and/or light
immunoglobulin transgenes. Specific segments of such
transgenes corresponding to functionally rearranged VDJ or VJ
segments, contain immunoglobulin DNA sequences which are also

clearly distinguishable from the endogenous immunoglobulin gene segments in the mouse.

Such differences in DNA sequence are also reflected in the amino acid sequence encoded by such human immunoglobulin transgenes as compared to those encoded by mouse B-cells. Thus, human immunoglobulin amino acid sequences may be detected in the transgenic non-human animals of the invention with antibodies specific for immunoglobulin epitopes encoded by human immunoglobulin gene segments.

Transgenic B-cells containing unrearranged transgenes from human or other species functionally recombine the appropriate gene segments to form functionally rearranged light and heavy chain variable regions. It will be readily apparent that the antibody encoded by such rearranged transgenes has a DNA and/or amino acid sequence which is heterologous to that normally encountered in the nonhuman animal used to practice the invention.

Unrearranged Transgenes

As used herein, an "unrearranged immunoglobulin heavy chain transgene" comprises DNA encoding at least one variable gene segment, one diversity gene segment, one joining gene segment and one constant region gene segment. Each of the gene segments of said heavy chain transgene are derived from, or has a sequence corresponding to, DNA encoding immunoglobulin heavy chain gene segments from a species not consisting of the non-human animal into which said transgene is introduced. Similarly, as used herein, an "unrearranged immunoglobulin light chain transgene" comprises DNA encoding at least one variable gene segment, one joining gene segment and at least one constant region gene segment wherein each gene segment of said light chain transgene is derived from, or has a sequence corresponding to, DNA encoding immunoglobulin light chain gene segments from a species not consisting of the non-human animal into which said light chain transgene is introduced.

Such heavy and light chain transgenes in this aspect of the invention contain the above-identified gene segments in

31

an unrearranged form. Thus, interposed between the V, D and J
segments in the heavy chain transgene and between the V and J
segments on the light chain transgene are appropriate
recombination signal sequences (RSS's). In addition, such

5 transgenes also include appropriate RNA splicing signals to
join a constant region gene segment with the VJ or VDJ
rearranged variable region.

In order to facilitate isotype switching within a
heavy chain transgene containing more than one C region gene

10 segment, e.g. $C\mu$ and $C\gamma1$ from the human genome, as explained
below "switch regions" are incorporated upstream from each of
the constant region gene segments and downstream from the
variable region gene segments to permit recombination between
such constant regions to allow for immunoglobulin class

15 switching, e.g. from IgM to IgG. Such heavy and light
immunoglobulin transgenes also contain transcription control
sequences including promoter regions situated upstream from
the variable region gene segments which typically contain TATA
motifs. A promoter region can be defined approximately as a

20 DNA sequence that, when operably linked to a downstream
sequence, can produce transcription of the downstream
sequence. Promoters may require the presence of additional
linked cis-acting sequences in order to produce efficient
transcription. In addition, other sequences that participate

25 in the transcription of sterile transcripts are preferably
included. Examples of sequences that participate in
expression of sterile transcripts can be found in the
published literature, including Rothman et al., Intl. Immunol.
2:621-627 (1990); Reid et al., Proc. Natl. Acad. Sci. USA

30 86:840-844 (1989); Stavnezer et al., Proc. Natl. Acad. Sci.
USA 85:7704-7708 (1988); and Mills et al., Nucl. Acids Res.
18:7305-7316 (1991), each of which is incorporated herein by
reference. These sequences typically include about at least
50 bp immediately upstream of a switch region, preferably

35 about at least 200 bp upstream of a switch region; and more
preferably about at least 200-1000 bp or more upstream of a
switch region. Suitable sequences occur immediately upstream
of the human $S_{\gamma1}$, $S_{\gamma2}$, $S_{\gamma3}$, $S_{\gamma4}$, $S_{\alpha1}$, $S_{\alpha2}$, and S_{ϵ} switch

32

regions, although the sequences immediately upstream of the human $S_{\gamma 1}$, and $S_{\gamma 3}$ switch regions are preferable. In particular, interferon (IFN) inducible transcriptional regulatory elements, such as IFN-inducible enhancers, are
5   preferably included immediately upstream of transgene switch sequences.

In addition to promoters, other regulatory sequences which function primarily in B-lineage cells are used. Thus, for example, a light chain enhancer sequence situated
10 preferably between the J and constant region gene segments on the light chain transgene is used to enhance transgene expression, thereby facilitating allelic exclusion. In the case of the heavy chain transgene, regulatory enhancers and also employed. Such regulatory sequences are used to maximize
15 the transcription and translation of the transgene so as to induce allelic exclusion and to provide relatively high levels of transgene expression.

Although the foregoing promoter and enhancer regulatory control sequences have been generically described,
20 such regulatory sequences may be heterologous to the nonhuman animal being derived from the genomic DNA from which the heterologous transgene immunoglobulin gene segments are obtained. Alternately, such regulatory gene segments are derived from the corresponding regulatory sequences in the
25 genome of the non-human animal, or closely related species, which contains the heavy and light transgene.

In the preferred embodiments, gene segments are derived from human beings. The transgenic non-human animals harboring such heavy and light transgenes are capable of
30 mounting an Ig-mediated immune response to a specific antigen administered to such an animal. B-cells are produced within such an animal which are capable of producing heterologous human antibody. After immortalization, and the selection for an appropriate monoclonal antibody (Mab), e.g. a hybridoma, a
35 source of therapeutic human monoclonal antibody is provided. Such human Mabs have significantly reduced immunogenicity when therapeutically administered to humans.

33

Although the preferred embodiments disclose the
construction of heavy and light transgenes containing human
gene segments, the invention is not so limited. In this
regard, it is to be understood that the teachings described
5   herein may be readily adapted to utilize immunoglobulin gene
segments from a species other than human beings. For example,
in addition to the therapeutic treatment of humans with the
antibodies of the invention, therapeutic antibodies encoded by
appropriate gene segments may be utilized to generate
10 monoclonal antibodies for use in the veterinary sciences.

Rearranged Transgenes
        In an alternative embodiment, transgenic nonhuman
animals contain functionally at least one rearranged
15 heterologous heavy chain immunoglobulin transgene in the
germline of the transgenic animal. Such animals contain
primary repertoire B-cells that express such rearranged heavy
transgenes. Such B-cells preferably are capable of undergoing
somatic mutation when contacted with an antigen to form a
20 heterologous antibody having high affinity and specificity for
the antigen. Said rearranged transgenes will contain at least
two C_H genes and the associated sequences required for isotype
switching.
        The invention also includes transgenic animals
25 containing germ line cells having heavy and light transgenes
wherein one of the said transgenes contains rearranged gene
segments with the other containing unrearranged gene segments.
In such animals, the heavy chain transgenes shall have at
least two C_H genes and the associated sequences required for
30 isotype switching.
        The invention further includes methods for
generating a synthetic variable region gene segment repertoire
to be used in the transgenes of the invention. The method
comprises generating a population of immunoglobulin V segment
35 DNAs wherein each of the V segment DNAs encodes an
immunoglobulin V segment and contains at each end a cleavage
recognition site of a restriction endonuclease. The
population of immunoglobulin V segment DNAs is thereafter

34

concatenated to form the synthetic immunoglobulin V segment repertoire. Such synthetic variable region heavy chain transgenes shall have at least two C_H genes and the associated sequences required for isotype switching.

5

Isotype Switching

In the development of a B lymphocyte, the cell initially produces IgM with a binding specificity determined by the productively rearranged V_H and V_L regions.
10   Subsequently, each B cell and its progeny cells synthesize antibodies with the same L and H chain V regions, but they may switch the isotype of the H chain.

The use of μ or δ constant regions is largely determined by alternate splicing, permitting IgM and IgD to be
15   coexpressed in a single cell. The other heavy chain isotypes (γ, α, and ϵ) are only expressed natively after a gene rearrangement event deletes the $C\mu$ and $C\delta$ exons. This gene rearrangement process, termed isotype switching, typically occurs by recombination between so called switch segments
20   located immediately upstream of each heavy chain gene (except δ). The individual switch segments are between 2 and 10 kb in length, and consist primarily of short repeated sequences. The exact point of recombination differs for individual class switching events. Investigations which have used solution
25   hybridization kinetics or Southern blotting with cDNA-derived C_H probes have confirmed that switching can be associated with loss of C_H sequences from the cell.

The switch (S) region of the μ gene, S_μ, is located about 1 to 2 kb 5' to the coding sequence and is composed of
30   numerous tandem repeats of sequences of the form $(GAGCT)_n(GGGGT)$, where n is usually 2 to 5 but can range as high as 17. (See T. Nikaido et al. Nature 292:845-848 (1981))

Similar internally repetitive switch sequences spanning several kilobases have been found 5' of the other C_H
35   genes. The Sα region has been sequenced and found to consist of tandemly repeated 80-bp homology units, whereas $S_{\gamma 2a}$, $S_{\gamma 2b}$, and $S_{\gamma 3}$ all contain repeated 49-bp homology units very similar to each other. (See, P. Szurek et al., J. Immunol 135:620-626

35

(1985) and T. Nikaido et al., J. Biol. Chem. 257:7322-7329
(1982), which are incorporated herein by reference.) All the
sequenced S regions include numerous occurrences of the
pentamers GAGCT and GGGGT that are the basic repeated elements
5   of the S_μ gene (T. Nikaido et al., J. Biol. Chem. 257:7322-
7329 (1982) which is incorporated herein by reference); in the
other S regions these pentamers are not precisely tandemly
repeated as in S_μ, but instead are embedded in larger repeat
units. The $S_{\gamma 1}$ region has an additional higher-order
10 structure: two direct repeat sequences flank each of two
clusters of 49-bp tandem repeats. (See M. R. Mowatt et al.,
J. Immunol. 136:2674-2683 (1986), which is incorporated herein
by reference).
        Switch regions of human H chain genes have been
15 found to be very similar to their mouse homologs. Indeed,
similarity between pairs of human and mouse clones 5' to the
C_H genes has been found to be confined to the S regions, a fact
that confirms the biological significance of these regions.
        A switch recombination between μ and α genes
20 produces a composite S_μ-S_α sequence. Typically, there is no
specific site, either in S_μ or in any other S region, where
the recombination always occurs.
        Generally, unlike the enzymatic machinery of V-J
recombination, the switch machinery can apparently accommodate
25 different alignments of the repeated homologous regions of
germline S precursors and then join the sequences at different
positions within the alignment. (See, T. H. Rabbits et al.,
Nucleic Acids Res. 9:4509-4524 (1981) and J. Ravetch et al.,
Proc. Natl. Acad. Sci. USA 77:6734-6738 (1980), which are
30 incorporated herein by reference.)
        The exact details of the mechanism(s) of selective
activation of switching to a particular isotype are unknown.
Although exogenous influences such as lymphokines and
cytokines might upregulate isotype-specific recombinases, it
35 is also possible that the same enzymatic machinery catalyzes
switches to all isotypes and that specificity lies in
targeting this machinery to specific switch regions.

36

The T-cell-derived lymphokines IL-4 and IFN$_\gamma$ have been shown to specifically promote the expression of certain isotypes: IL-4 decreases IgM, IgG2a, IgG2b, and IgG3 expression and increases IgE and IgG1 expression; while IFN$_\gamma$
5   selectively stimulates IgG2a expression and antagonizes the IL-4-induced increase in IgE and IgG1 expression (Coffman et al., *J. Immunol.* 136:949-954 (1986) and Snapper et al., *Science* 236:944-947 (1987), which are incorporated herein by reference). A combination of IL-4 and IL-5 promotes IgA
10 expression (Coffman et al., *J. Immunol.* 139:3685-3690 (1987), which is incorporated herein by reference).

Most of the experiments implicating T-cell effects on switching have not ruled out the possibility that the observed increase in cells with particular switch
15 recombinations might reflect selection of preswitched or precommitted cells; but the most likely explanation is that the lymphokines actually promote switch recombination.

Induction of class switching appears to be associated with sterile transcripts that initiate upstream of
20 the switch segments (Lutzker et al., *Mol. Cell. Biol.* 8:1849 (1988); Stavnezer et al., *Proc. Natl. Acad. Sci. USA* 85:7704 (1988); Esser and Radbruch, *EMBO J.* 8:483 (1989); Berton et al., *Proc. Natl. Acad. Sci. USA* 86:2829 (1989); Rothman et al., *Int. Immunol.* 2:621 (1990), each of which is incorporated
25 by reference). For example, the observed induction of the γ1 sterile transcript by IL-4 and inhibition by IFN-γ correlates with the observation that IL-4 promotes class switching to γ1 in B-cells in culture, while IFN-γ inhibits γ1 expression. Therefore, the inclusion of regulatory sequences that affect
30 the transcription of sterile transcripts may also affect the rate of isotype switching. For example, increasing the transcription of a particular sterile transcript typically can be expected to enhance the frequency of isotype switch recombination involving adjacent switch sequences.

35 For these reasons, it is preferable that transgenes incorporate transcriptional regulatory sequences within about 1-2 kb upstream of each switch region that is to be utilized for isotype switching. These transcriptional regulatory

sequences preferably include a promoter and an enhancer element, and more preferably include the 5' flanking (i.e., upstream) region that is naturally associated (i.e., occurs in germline configuration) with a switch region. This 5'
5   flanking region is typically about at least 50 nucleotides in length, preferably about at least 200 nucleotides in length, and more preferably at least 500-1000 nucleotides.

Although a 5' flanking sequence from one switch region can be operably linked to a different switch region for
10 transgene construction (e.g., the 5' flanking sequence from the human $S_{\gamma 1}$ switch can be grafted immediately upstream of the $S_{\alpha 1}$ switch), in some embodiments it is preferred that each switch region incorporated in the transgene construct have the 5' flanking region that occurs immediately upstream in the
15 naturally occurring germline configuration.

The Transgenic Primary Repertoire

A.   The Human Immunoglobulin Loci

An important requirement for transgene function is
20 the generation of a primary antibody repertoire that is diverse enough to trigger a secondary immune response for a wide range of antigens. The rearranged heavy chain gene consists of a signal peptide exon, a variable region exon and a tandem array of multi-domain constant region regions, each
25 of which is encoded by several exons. Each of the constant region genes encode the constant portion of a different class of immunoglobulins. During B-cell development, V region proximal constant regions are deleted leading to the expression of new heavy chain classes. For each heavy chain
30 class, alternative patterns of RNA splicing give rise to both transmembrane and secreted immunoglobulins.

The human heavy chain locus consists of approximately 200 V gene segments spanning 2 Mb, approximately 30 D gene segments spanning about 40 kb, six J segments
35 clustered within a 3 kb span, and nine constant region gene segments spread out over approximately 300 kb. The entire locus spans approximately 2.5 Mb of the distal portion of the long arm of chromosome 14.

B.   Gene Fragment Transgenes

        1. Heavy Chain Transgene

            In a preferred embodiment, immunoglobulin heavy and
light chain transgenes comprise unrearranged genomic DNA from

5 humans. In the case of the heavy chain, a preferred transgene
comprises a NotI fragment having a length between 670 to 830
kb. The length of this fragment is ambiguous because the 3'
restriction site has not been accurately mapped. It is known,
however, to reside between the $\alpha 1$ and $\psi \alpha$ gene segments. This

10 fragment contains members of all six of the known V_H families,
the D and J gene segments, as well as the μ, δ, $\gamma 3$, $\gamma 1$ and $\alpha 1$
constant regions (Berman et al., EMBO J. 7:727-738 (1988),
which is incorporated herein by reference). A transgenic
mouse line containing this transgene correctly expresses a

15 heavy chain class required for B-cell development (IgM) and at
least one switched heavy chain class (IgG$_1$), in conjunction
with a sufficiently large repertoire of variable regions to
trigger a secondary response for most antigens.

20   2. Light Chain Transgene

            A genomic fragment containing all of the necessary
gene segments and regulatory sequences from a human light
chain locus may be similarly constructed. Such transgenes are
constructed as described in the Examples.

25
C.   Transgenes Generated Intracellularly
      by In Vivo Recombination

            It is not necessary to isolate the all or part of
the heavy chain locus on a single DNA fragment. Thus, for

30 example, the 670-830 kb NotI fragment from the human
immunoglobulin heavy chain locus may be formed in vivo in the
non-human animal during transgenesis. Such in vivo transgene
construction is produced by introducing two or more
overlapping DNA fragments into an embryonic nucleus of the

35 non-human animal. The overlapping portions of the DNA
fragments have DNA sequences which are substantially
homologous. Upon exposure to the recombinases contained
within the embryonic nucleus, the overlapping DNA fragments

homologously recombined in proper orientation to form the
670-830 kb NotI heavy chain fragment.

        In vivo transgene construction can be used to form
any number of immunoglobulin transgenes which because of their
size are otherwise difficult, or impossible, to make or
manipulate by present technology. Thus, _in vivo_ transgene
construction is useful to generate immunoglobulin transgenes
which are larger than DNA fragments which may be manipulated
by YAC vectors (Murray and Szostak, _Nature_ 305:189-193
(1983)). Such _in vivo_ transgene construction may be used to
introduce into a non-human animal substantially the entire
immunoglobulin loci from a species not consisting of the
transgenic non-human animal.

        In addition to forming genomic immunoglobulin
transgenes, _in vivo_ homologous recombination may also be
utilized to form "mini-locus" transgenes as described in the
Examples.

        In the preferred embodiments utilizing _in vivo_
transgene construction, each overlapping DNA fragment
preferably has an overlapping substantially homologous DNA
sequence between the end portion of one DNA fragment and the
end portion of a second DNA fragment. Such overlapping
portions of the DNA fragments preferably comprise about 500 bp
to about 2000 bp, most preferably 1.0 kb to 2.0 kb.
Homologous recombination of overlapping DNA fragments to form
transgenes _in vivo_ is further described in commonly assigned
PCT Publication No. WO 92/03917 entitled "Homologous
Recombination in Mammalian Cells" published March 19, 1992.

D.   **Minilocus Transgenes**

        As used herein, the term "immunoglobulin minilocus"
refers to a DNA sequence (which may be within a longer
sequence), usually of less than about 150 kb, typically
between about 25 and 100 kb, containing at least one each of
the following: a functional variable (V) gene segment, a
functional joining (J) region segment, at least one functional
constant (C) region gene segment, and--if it is a heavy chain
minilocus--a functional diversity (D) region segment, such

that said DNA sequence contains at least one substantial
discontinuity (e.g., a deletion, usually of at least about 2
to 5 kb, preferably 10-25 kb or more, relative to the
homologous genomic DNA sequence). A light chain minilocus
5   transgene will be at least 25 kb in length, typically 50 to 60
kb. A heavy chain transgene will typically be about 70 to 80
kb in length, preferably at least about 60 kb with two
constant regions operably linked to switch regions.
Furthermore, the individual elements of the minilocus are
10 preferably in the germline configuration and capable of
undergoing gene rearrangement in the pre-B cell of a
transgenic animal so as to express functional antibody
molecules with diverse antigen specificities encoded entirely
by the elements of the minilocus. Further, a heavy chain
15 minilocus comprising at least two C_H genes and the requisite
switching sequences is typically capable of undergoing isotype
switching, so that functional antibody molecules of different
immunoglobulin classes will be generated. Such isotype
switching may occur in vivo in B-cells residing within the
20 transgenic nonhuman animal, or may occur in cultured cells of
the B-cell lineage which have been explanted from the
transgenic nonhuman animal.

        In an alternate preferred embodiment, immunoglobulin
heavy chain transgenes comprise one or more of each of the V_H,
25 D, and J_H gene segments and two or more of the C_H genes. At
least one of each appropriate type gene segment is
incorporated into the minilocus transgene. With regard to the
C_H segments for the heavy chain transgene, it is preferred
that the transgene contain at least one μ gene segment and at
30 least one other constant region gene segment, more preferably
a γ gene segment, and most preferably $\gamma3$ or $\gamma1$. This
preference is to allow for class switching between IgM and IgG
forms of the encoded immunoglobulin and the production of a
secretable form of high affinity non-IgM immunoglobulin.
35 Other constant region gene segments may also be used such as
those which encode for the production of IgD, IgA and IgE.

        Those skilled in the art will also construct
transgenes wherein the order of occurrence of heavy chain C_H

41

genes will be different from the naturally-occurring spatial
order found in the germline of the species serving as the
donor of the C_H genes.

Additionally, those skilled in the art can select C_H
5 genes from more than one individual of a species (e.g.,
allogeneic C_H genes) and incorporate said genes in the
transgene as supernumerary C_H genes capable of undergoing
isotype switching; the resultant transgenic nonhuman animal
may then, in some embodiments, make antibodies of various
10 classes including all of the allotypes represented in the
species from which the transgene C_H genes were obtained.

Still further, those skilled in the art can select
C_H genes from different species to incorporate into the
transgene. Functional switch sequences are included with each
15 C_H gene, although the switch sequences used are not
necessarily those which occur naturally adjacent to the C_H
gene. Interspecies C_H gene combinations will produce a
transgenic nonhuman animal which may produce antibodies of
various classes corresponding to C_H genes from various
20 species. Transgenic nonhuman animals containing interspecies
C_H transgenes may serve as the source of B-cells for
constructing hybridomas to produce monoclonals for veterinary
uses.

The heavy chain J region segments in the human
25 comprise six functional J segments and three pseudo genes
clustered in a 3 kb stretch of DNA. Given its relatively
compact size and the ability to isolate these segments
together with the μ gene and the 5' portion of the δ gene on a
single 23 kb SFiI/SpeI fragment (Sado et al., Biochem.
30 Biophys. Res. Comm. 154:264271 (1988), which is incorporated
herein by reference); it is preferred that all of the J region
gene segments be used in the mini-locus construct. Since this
fragment spans the region between the μ and δ genes, it is
likely to contain all of the 3' cis-linked regulatory elements
35 required for μ expression. Furthermore, because this fragment
includes the entire J region, it contains the heavy chain
enhancer and the μ switch region (Mills et al., Nature 306:809
(1983); Yancopoulos and Alt, Ann. Rev. Immunol. 4:339-368

42

(1986), which are incorporated herein by reference). It also
contains the transcription start sites which trigger VDJ
joining to form primary repertoire B-cells (Yancopoulos and
Alt, Cell 40:271-281 (1985), which is incorporated herein by
5   reference). Alternatively, a 36 kb BssHII/SpeI1 fragment,
which includes part on the D region, may be used in place of
the 23 kb SfiI/SpeI1 fragment. The use of such a fragment
increases the amount of 5' flanking sequence to facilitate
efficient D-to-J joining.

10   The human D region consists of 4 or 5 homologous 9
kb subregions, linked in tandem (Siebenlist, et al. (1981),
Nature, 294, 631-635). Each subregion contains up to 10
individual D segments. Some of these segments have been
mapped and are shown in Fig. 4. Two different strategies are
15 used to generate a mini-locus D region. The first strategy
involves using only those D segments located in a short
contiguous stretch of DNA that includes one or two of the
repeated D subregions. A candidate is a single 15 kb fragment
that contains 12 individual D segments. This piece of DNA
20 consists of 2 contiguous EcoRI fragments and has been
completely sequenced (Ichihara, et al. (1988), EMBO J., 7,
4141-4150). Twelve D segments should be sufficient for a
primary repertoire. However, given the dispersed nature of
the D region, an alternative strategy is to ligate together
25 several non-contiguous D-segment containing fragments, to
produce a smaller piece of DNA with a greater number of
segments. Additional D-segment genes can be identified, for
example, by the presence of characteristic flanking nonamer
and heptamer sequences, supra, and by reference to the
30 literature.
          At least one, and preferably more than one V gene
segment is used to construct the heavy chain minilocus
transgene. Rearranged or unrearranged V segments with or
without flanking sequences can be isolated as described PCT
35 Publication No. WO 92/03918, published March 19, 1992,
entitled "Transgenic Non-Human Animals Capable of Producing
Heterologous Antibodies."

43

Rearranged or unrearranged V segments, D segments, J segments, and C genes, with or without flanking sequences, can be isolated as described in PCT Publication No. WO 92/03918, published March 19, 1992.

5   A minilocus light chain transgene may be similarly constructed from the human λ or κ immunoglobulin locus. Thus, for example, an immunoglobulin heavy chain minilocus transgene construct, e.g., of about 75 kb, encoding V, D, J and constant region sequences can be formed from a plurality

10   of DNA fragments, with each sequence being substantially homologous to human gene sequences. Preferably, the sequences are operably linked to transcription regulatory sequences and are capable of undergoing rearrangement. With two or more appropriately placed constant region sequences (e.g., μ and γ)

15   and switch regions, switch recombination also occurs. An exemplary light chain transgene construct can be formed similarly from a plurality of DNA fragments, substantially homologous to human DNA and capable of undergoing rearrangement.

20

E.   Transgene Constructs Capable of Isotype Switching

Ideally, transgene constructs that are intended to undergo class switching should include all of the cis-acting sequences necessary to regulate sterile transcripts.

25   Naturally occurring switch regions and upstream promoters and regulatory sequences (e.g., IFN-inducible elements) are preferred cis-acting sequences that are included in transgene constructs capable of isotype switching. About at least 50 basepairs, preferably about at least 200 basepairs, and more

30   preferably at least 500 to 1000 basepairs or more of sequence immediately upstream of a switch region, preferably a human $\gamma 1$ switch region, should be operably linked to a switch sequence, preferably a human $\gamma 1$ switch sequence. Further, switch regions can be linked upstream of (and adjacent to) C_H genes

35   that do not naturally occur next to the particular switch region. For example, but not for limitation, a human γ_1 switch region may be linked upstream from a human α_2 C_H gene, or a murine γ_1 switch may be linked to a human C_H gene.

44

An alternative method for obtaining non-classical
isotype switching (e.g., δ-associated deletion) in transgenic
mice involves the inclusion of the 400 bp direct repeat
sequences (σμ and εμ) that flank the human μ gene (Yasui et

5   al., Eur. J. Immunol. 19:1399 (1989)). Homologous
recombination between these two sequences deletes the μ gene
in IgD-only B-cells. Heavy chain transgenes can be
represented by the following formulaic description:

10   $(V_H)_x - (D)_y - (J_H)_z - (S_D)_m - (C_1)_n - [(T) - (S_A)_p - (C_2)]_q$

where:
    V_H is a heavy chain variable region gene segment,
    D is a heavy chain D (diversity) region gene segment,
15   J_H is a heavy chain J (joining) region gene segment,
    S_D is a donor region segment capable of participating in
    a recombination event with the S_a acceptor region
    segments such that isotype switching occurs,
    C_1 is a heavy chain constant region gene segment encoding
20   an isotype utilized in for B cell development (e.g.,
         μ or δ),
    T is a cis-acting transcriptional regulatory region
         segment containing at least a promoter,
    S_A is an acceptor region segment capable of participating
25   in a recombination event with selected S_D donor
         region segments, such that isotype switching occurs,
    C_2 is a heavy chain constant region gene segment encoding
         an isotype other than μ (e.g., γ_1, γ_2, γ_3, γ_4, α_1,
         α_2, ε).
30   x, y, z, m, n, p, and q are integers. x is 1-100, n is
         0-10, y is 1-50, p is 1-10, z is 1-50, q is 0-50, m
         is 0-10. Typically, when the transgene is capable
         of isotype switching, q must be at least 1, m is at
         least 1, n is at least 1, and m is greater than or
35   equal to n.

45

V_H, D, J_H, S_D, C_1, T, S_A, and C_Z segments may be
selected from various species, preferably mammalian species,
and more preferably from human and murine germline DNA.

V_H segments may be selected from various species,
5   but are preferably selected from V_H segments that occur
naturally in the human germline, such as V_{H251}. Typically
about 2 V_H gene segments are included, preferably about 4 V_H
segments are included, and most preferably at least about 10
V_H segments are included.

10   At least one D segment is typically included,
although at least 10 D segments are preferably included, and
some embodiments include more than ten D segments. Some
preferred embodiments include human D segments.

Typically at least one J_H segment is incorporated in
15 the transgene, although it is preferable to include about six
J_H segments, and some preferred embodiments include more than
about six J_H segments. Some preferred embodiments include
human J_H segments, and further preferred embodiments include
six human J_H segments and no nonhuman J_H segments.

20   S_D segments are donor regions capable of
participating in recombination events with the S_A segment of
the transgene. For classical isotype switching, S_D and S_A are
switch regions such as S_μ, $S_{\gamma 1}$, $S_{\gamma 2}$, $S_{\gamma 3}$, $S_{\gamma 4}$, S_α, $S_{\alpha 2}$, and S_ϵ.
Preferably the switch regions are murine or human, more
25 preferably S_D is a human or murine S_μ and S_A is a human or
murine $S_{\gamma 1}$. For nonclassical isotype switching (δ-associated
deletion), S_D and S_A are preferably the 400 basepair direct
repeat sequences that flank the human μ gene.

C_1 segments are typically μ or δ genes, preferably a
30 μ gene, and more preferably a human or murine μ gene.

T segments typically include 5' flanking sequences
that are adjacent to naturally occurring (i.e., germline)
switch regions. T segments typically at least about at least
50 nucleotides in length, preferably about at least 200
35 nucleotides in length, and more preferably at least 500-1000
nucleotides in length. Preferably T segments are 5' flanking
sequences that occur immediately upstream of human or murine
switch regions in a germline configuration. It is also

evident to those of skill in the art that T segments may
comprise cis-acting transcriptional regulatory sequences that
do not occur naturally in an animal germline (e.g., viral
enhancers and promoters such as those found in SV40,
5 adenovirus, and other viruses that infect eukaryotic cells).

C_2 segments are typically a γ_1, γ_2, γ_3, γ_4, α_1, α_2,
or ϵ C_H gene, preferably a human C^H gene of these isotypes, and
more preferably a human γ_1 or γ_3 gene. Murine γ_{2a} and γ_{2b} may
also be used, as may downstream (i.e., switched) isotype genes
10 form various species. Where the heavy chain transgene
contains an immunoglobulin heavy chain minilocus, the total
length of the transgene will be typically 150 kilo basepairs
or less.

In general, the transgene will be other than a
15 native heavy chain Ig locus. Thus, for example, deletion of
unnecessary regions or substitutions with corresponding
regions from other species will be present.

F.   Methods for Determining Functional
20   Isotype Switching in Ig Transgenes

The occurrence of isotype switching in a transgenic
nonhuman animal may be identified by any method known to those
in the art. Preferred embodiments include the following,
25 employed either singly or in combination:

1. detection of mRNA transcripts that contain a sequence
homologous to at least one transgene downstream C_H gene other
than δ and an adjacent sequence homologous to a transgene V_H-
D_H-J_H rearranged gene; such detection may be by Northern
30 hybridization, S_1 nuclease protection assays, PCR
amplification, cDNA cloning, or other methods;

2. detection in the serum of the transgenic animal, or in
supernatants of cultures of hybridoma cells made from B-cells
of the transgenic animal, of immunoglobulin proteins encoded
35 by downstream C_H genes, where such proteins can also be shown
by immunochemical methods to comprise a functional variable
region;

3. detection, in DNA from B-cells of the transgenic
animal or in genomic DNA from hybridoma cells, of DNA

47

rearrangements consistent with the occurrence of isotype switching in the transgene, such detection may be accomplished by Southern blot hybridization, PCR amplification, genomic cloning, or other method; or

5   4. identification of other indicia of isotype switching, such as production of sterile transcripts, production of characteristic enzymes involved in switching (e.g., "switch recombinase"), or other manifestations that may be detected, measured, or observed by contemporary techniques.

10   Because each transgenic line may represent a different site of integration of the transgene, and a potentially different tandem array of transgene inserts, and because each different configuration of transgene and flanking DNA sequences can affect gene expression, it is preferable to

15 identify and use lines of mice that express high levels of human immunoglobulins, particularly of the IgG isotype, and contain the least number of copies of the transgene. Single copy transgenics minimize the potential problem of incomplete allelic expression. Transgenes are typically integrated into

20 host chromosomal DNA, most usually into germline DNA and propagated by subsequent breeding of germline transgenic breeding stock animals. However, other vectors and transgenic methods known in the present art or subsequently developed may be substituted as appropriate and as desired by a

25 practitioner.

G.   Functional Disruption of
     Endogenous Immunoglobulin Loci

30   The expression of successfully rearranged immunoglobulin heavy and light transgenes is expected to have a dominant effect by suppressing the rearrangement of the endogenous immunoglobulin genes in the transgenic nonhuman animal. However, another way to generate a nonhuman that is

35 devoid of endogenous antibodies is by mutating the endogenous immunoglobulin loci. Using embryonic stem cell technology and homologous recombination, the endogenous immunoglobulin repertoire can be readily eliminated. The following describes the functional description of the mouse immunoglobulin loci.

48

The vectors and methods disclosed, however, can be readily adapted for use in other non-human animals.

Briefly, this technology involves the inactivation of a gene, by homologous recombination, in a pluripotent cell
5   line that is capable of differentiating into germ cell tissue. A DNA construct that contains an altered, copy of a mouse immunoglobulin gene is introduced into the nuclei of embryonic stem cells. In a portion of the cells, the introduced DNA recombines with the endogenous copy of the mouse gene,
10 replacing it with the altered copy. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is reimplanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells entirely derived from the mutant cell line. Therefore,
15 by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic lesion (reviewed by Capecchi (1989), Science, 244, 1288-1292).

Because the mouse λ locus contributes to only 5% of the immunoglobulins, inactivation of the heavy chain and/or
20 κ-light chain loci is sufficient. There are three ways to disrupt each of these loci, deletion of the J region, deletion of the J-C intron enhancer, and disruption of constant region coding sequences by the introduction of a stop codon. The last option is the most straightforward, in terms of DNA
25 construct design. Elimination of the μ gene disrupts B-cell maturation thereby preventing class switching to any of the functional heavy chain segments. The strategy for knocking out these loci is outlined below.

To disrupt the mouse μ and κ genes, targeting
30 vectors are used based on the design employed by Jaenisch and co-workers (Zijlstra, et al. (1989), Nature, 342, 435-438) for the successful disruption of the mouse β2-microglobulin gene. The neomycin resistance gene (neo), from the plasmid pMCIneo is inserted into the coding region of the target gene. The
35 pMCIneo insert uses a hybrid viral promoter/enhancer sequence to drive neo expression. This promoter is active in embryonic stem cells. Therefore, neo can be used as a selectable marker for integration of the knock-out construct. The HSV thymidine

kinase (tk) gene is added to the end of the construct as a
negative selection marker against random insertion events
(Zijlstra, et al., supra.).

5   A preferred strategy for disrupting the heavy chain
locus is the elimination of the J region. This region is
fairly compact in the mouse, spanning only 1.3 kb. To
construct a gene targeting vector, a 15 kb KpnI fragment
containing all of the secreted A constant region exons from
mouse genomic library is isolated. The 1.3 kb J region is
10 replaced with the 1.1 kb insert from pMCIneo. The HSV tk gene
is then added to the 5' end of the KpnI fragment. Correct
integration of this construct, via homologous recombination,
will result in the replacement of the mouse J_H region with the
neo gene. Recombinants are screened by PCR, using a primer
15 based on the neo gene and a primer homologous to mouse
sequences 5' of the KpnI site in the D region.

          Alternatively, the heavy-chain locus is knocked out
by disrupting the coding region of the μ gene. This approach
involves the same 15 kb KpnI fragment used in the previous
20 approach. The 1.1 kb insert from pMCIneo is inserted at a
unique BamHI site in exon II, and the HSV tk gene added to the
3' KpnI end. Double crossover events on either side of the
neo insert, that eliminate the tk gene, are then selected for.
These are detected from pools of selected clones by PCR
25 amplification. One of the PCR primers is derived from neo
sequences and the other from mouse sequences outside of the
targeting vector. The functional disruption of the mouse
immunoglobulin loci is presented in the Examples.

30

G.   Suppressing Expression of
      Endogenous Immunoglobulin Loci
          In addition to functional disruption of endogenous
Ig loci, an alternative method for preventing the expression
35 of an endogenous Ig locus is suppression. Suppression of
endogenous Ig genes may be accomplished with antisense RNA
produced from one or more integrated transgenes, by antisense
oligonucleotides, and/or by administration of antisera
specific for one or more endogenous Ig chains.

50

Antisense Polynucleotides

Antisense RNA transgenes can be employed to
partially or totally knock-out expression of specific genes
(Pepin et al. (1991) Nature 355: 725; Helene., C. and Toulme,
5 J. (1990) Biochimica Biophys. Acta 1049: 99; Stout, J. and
Caskey, T. (1990) Somat. Cell Mol. Genet. 16: 369; Munir et
al. (1990) Somat. Cell Mol. Genet. 16: 383, each of which is
incorporated herein by reference).

"Antisense polynucleotides" are polynucleotides
10 that: (1) are complementary to all or part of a reference
sequence, such as a sequence of an endogenous Ig C_H or C_L
region, and (2) which specifically hybridize to a
complementary target sequence, such as a chromosomal gene
locus or a Ig mRNA. Such complementary antisense
15 polynucleotides may include nucleotide substitutions,
additions, deletions, or transpositions, so long as specific
hybridization to the relevant target sequence is retained as a
functional property of the polynucleotide. Complementary
antisense polynucleotides include soluble antisense RNA or DNA
20 oligonucleotides which can hybridize specifically to
individual mRNA species and prevent transcription and/or RNA
processing of the mRNA species and/or translation of the
encoded polypeptide (Ching et al., Proc. Natl. Acad. Sci.
U.S.A. 86:10006-10010 (1989); Broder et al., Ann. Int. Med.
25 113:604-618 (1990); Loreau et al., FEBS Letters 274:53-56
(1990); Holcenberg et al., WO91/11535; WO91/09865; WO91/04753;
WO90/13641; and EP 386563, each of which is incorporated
herein by reference). An antisense sequence is a
polynucleotide sequence that is complementary to at least one
30 immunoglobulin gene sequence of at least about 15 contiguous
nucleotides in length, typically at least 20 to 30 nucleotides
in length, and preferably more than about 30 nucleotides in
length. However, in some embodiments, antisense sequences may
have substitutions, additions, or deletions as compared to the
35 complementary immunoglobulin gene sequence, so long as
specific hybridization is retained as a property of the
antisense polynucleotide. Generally, an antisense sequence is
complementary to an endogenous immunoglobulin gene sequence

51

that encodes, or has the potential to encode after DNA
rearrangement, an immunoglobulin chain. In some cases, sense
sequences corresponding to an immunoglobulin gene sequence may
function to suppress expression, particularly by interfering
5   with transcription.

The antisense polynucleotides therefore inhibit
production of the encoded polypeptide(s). In this regard,
antisense polynucleotides that inhibit transcription and/or
translation of one or more endogenous Ig loci can alter the
10   capacity and/or specificity of a non-human animal to produce
immunoglobulin chains encoded by endogenous Ig loci.

Antisense polynucleotides may be produced from a
heterologous expression cassette in a transfectant cell or
transgenic cell, such as a transgenic pluripotent
15   hematopoietic stem cell used to reconstitute all or part of
the hematopoietic stem cell population of an individual, or a
transgenic nonhuman animal. Alternatively, the antisense
polynucleotides may comprise soluble oligonucleotides that are
administered to the external milieu, either in culture medium
20   in vitro or in the circulatory system or interstitial fluid in
vivo. Soluble antisense polynucleotides present in the
external milieu have been shown to gain access to the
cytoplasm and inhibit translation of specific mRNA species. In
some embodiments the antisense polynucleotides comprise
25   methylphosphonate moieties, alternatively phosphorothiolates
or O-methylribonucleotides may be used, and chimeric
oligonucleotides may also be used (Dagle et al. (1990) Nucleic
Acids Res. 18: 4751). For some applications, antisense
oligonucleotides may comprise polyamide nucleic acids (Nielsen
30   et al. (1991) Science 254: 1497). For general methods
relating to antisense polynucleotides, see Antisense RNA and
DNA, (1988), D.A. Melton, Ed., Cold Spring Harbor Laboratory,
Cold Spring Harbor, NY).

Antisense polynucleotides complementary to one or
35   more sequences are employed to inhibit transcription, RNA
processing, and/or translation of the cognate mRNA species and
thereby effect a reduction in the amount of the respective
encoded polypeptide. Such antisense polynucleotides can

provide a therapeutic function by inhibiting the formation of one or more endogenous Ig chains in vivo.

Whether as soluble antisense oligonucleotides or as antisense RNA transcribed from an antisense transgene, the

5 antisense polynucleotides of this invention are selected so as to hybridize preferentially to endogenous Ig sequences at physiological conditions in vivo. Most typically, the selected antisense polynucleotides will not appreciably hybridize to heterologous Ig sequences encoded by a heavy or

10 light chain transgene of the invention (i.e., the antisense oligonucleotides will not inhibit transgene Ig expression by more than about 25 to 35 percent).

Antiserum Suppression

15 Partial or complete suppression of endogenous Ig chain expression can be produced by injecting mice with antisera against one or more endogenous Ig chains (Weiss et al. (1984) Proc. Natl. Acad. Sci. (U.S.A.) 81 211, which is incorporated herein by reference). Antisera are selected so

20 as to react specifically with one or more endogenous Ig chains but to have minimal or no cross-reactivity with heterologous Ig chains encoded by an Ig transgene of the invention. Thus, administration of selected antisera according to a schedule as typified by that of Weiss et al. op.cit. will suppress

25 endogenous Ig chain expression but permits expression of heterologous Ig chain(s) encoded by a transgene of the present invention.

Nucleic Acids

30 The nucleic acids, the term "substantial homology" indicates that two nucleic acids, or designated sequences thereof, when optimally aligned and compared, are identical, with appropriate nucleotide insertions or deletions, in at least about 80% of the nucleotides, usually at least about 90%

35 to 95%, and more preferably at least about 98 to 99.5% of the nucleotides. Alternatively, substantial homology exists when the segments will hybridize under selective hybridization conditions, to the complement of the strand. The nucleic

53

acids may be present in whole cells, in a cell lysate, or in a
partially purified or substantially pure form. A nucleic acid
is "isolated" or "rendered substantially pure" when purified
away from other cellular components or other contaminants,

5   e.g., other cellular nucleic acids or proteins, by standard
techniques, including alkaline/SDS treatment, CsCl banding,
column chromatography, agarose gel electrophoresis and others
well known in the art. See, F. Ausubel, et al., ed. Current
Protocols in Molecular Biology, Greene Publishing and Wiley-

10 Interscience, New York (1987).

        The nucleic acid compositions of the present
invention, while often in a native sequence (except for
modified restriction sites and the like), from either cDNA,
genomic or mixtures may be mutated, thereof in accordance with

15 standard techniques to provide gene sequences. For coding
sequences, these mutations, may affect amino acid sequence as
desired. In particular, DNA sequences substantially
homologous to or derived from native V, D, J, constant,
switches and other such sequences described herein are

20 contemplated (where "derived" indicates that a sequence is
identical or modified from another sequence).

        A nucleic acid is "operably linked" when it is
placed into a functional relationship with another nucleic
acid sequence. For instance, a promoter or enhancer is

25 operably linked to a coding sequence if it affects the
transcription of the sequence. With respect to transcription
regulatory sequences, operably linked means that the DNA
sequences being linked are contiguous and, where necessary to
join two protein coding regions, contiguous and in reading

30 frame. For switch sequences, operably linked indicates that
the sequences are capable of effecting switch recombination.

Specific Preferred Embodiments

        A preferred embodiment of the invention is an animal

35 containing at least one, typically 2-10, and sometimes 25-50
or more copies of the transgene described in Example 12 (e.g.,
pHC1 or pHC2) bred with an animal containing a single copy of
a light chain transgene described in Examples 5, 6, 8, or 14,

54

and the offspring bred with the J_H deleted animal described in
Example 10. Animals are bred to homozygosity for each of
these three traits. Such animals have the following genotype:
a single copy (per haploid set of chromosomes) of a human

5   heavy chain unrearranged mini-locus (described in Example 12),
a single copy (per haploid set of chromosomes) of a rearranged
human κ light chain construct (described in Example 14), and a
deletion at each endogenous mouse heavy chain locus that
removes all of the functional J_H segments (described in

10   Example 10). Such animals are bred with mice that are
homozygous for the deletion of the J_H segments (Examples 10)
to produce offspring that are homozygous for the J_H deletion
and hemizygous for the human heavy and light chain constructs.
The resultant animals are injected with antigens and used for

15   production of human monoclonal antibodies against these
antigens.
          B cells isolated from such an animal are
monospecific with regard to the human heavy and light chains
because they contain only a single copy of each gene.

20   Furthermore, they will be monospecific with regards to human
or mouse heavy chains because both endogenous mouse heavy
chain gene copies are nonfunctional by virtue of the deletion
spanning the J_H region introduced as described in Example 9
and 12. Furthermore, a substantial fraction of the B cells

25   will be monospecific with regards to the human or mouse light
chains because expression of the single copy of the rearranged
human κ light chain gene will allelically and isotypically
exclude the rearrangement of the endogenous mouse κ and λ
chain genes in a significant fraction of B-cells.

30   The transgenic mouse of the preferred embodiment
will exhibit immunoglobulin production with a significant
repertoire, ideally substantially similar to that of a native
mouse. Thus, for example, in embodiments where the endogenous
Ig genes have been inactivated, the total immunoglobulin

35   levels will range from about 0.1 to 10 mg/ml of serum,
preferably 0.5 to 5 mg/ml, ideally at least about 1.0 mg/ml.
When a transgene capable of effecting a switch to IgG from IgM
has been introduced into the transgenic mouse, the adult mouse

55

ratio of serum IgG to IgM is preferably about 10:1.   Of
course, the IgG to IgM ratio will be much lower in the
immature mouse.   In general, greater than about 10%,
preferably 40 to 80% of the spleen and lymph node B cells
express exclusively human IgG protein.

The repertoire will ideally approximate that shown
in a non-transgenic mouse, usually at least about 10% as high,
preferably 25 to 50% or more.   Generally, at least about a
thousand different immunoglobulins (ideally IgG), preferably
10^4 to 10^6 or more, will be produced, depending primarily on
the number of different V, J and D regions introduced into the
mouse genome.   These immunoglobulins will typically recognize
about one-half or more of highly antigenic proteins,
including, but not limited to: pigeon cytochrome C, chicken
lysozyme, pokeweed mitogen, bovine serum albumin, keyhole
limpit hemocyanin, influenza hemagglutinin, staphylococcus
protein A, sperm whale myoglobin, influenza neuraminidase, and
lambda repressor protein.   Some of the immunoglobulins will
exhibit an affinity for preselected antigens of at least about
$10^7 \underline{M}^{-1}$, preferably $10^8 \underline{M}^{-1}$ to $10^9 \underline{M}^{-1}$ or greater.

Thus, prior to rearrangement of a transgene
containing various heavy or light chain gene segments, such
gene segments may be readily identified, e.g. by hybridization
or DNA sequencing, as being from a species of organism other
than the transgenic animal.

Although the foregoing describes a preferred
embodiment of the transgenic animal of the invention, other
embodiments are defined by the disclosure herein and more
particularly by the transgenes described in the Examples.
Four categories of transgenic animal may be defined:

I.   Transgenic animals containing an unrearranged heavy
and rearranged light immunoglobulin transgene.

II.   Transgenic animals containing an unrearranged heavy
and unrearranged light immunoglobulin transgene

III. Transgenic animal containing rearranged heavy and an
unrearranged light immunoglobulin transgene, and

IV.   Transgenic animals containing rearranged heavy and
rearranged light immunoglobulin transgenes.

56

Of these categories of transgenic animal, the preferred order of preference is as follows II > I > III > IV where the endogenous light chain genes (or at least the κ gene) have been knocked out by homologous recombination (or

5   other method) and I > II > III > IV where the endogenous light chain genes have not been knocked out and must be dominated by allelic exclusion.

EXPERIMENTAL EXAMPLES

10 METHODS AND MATERIALS

       Transgenic mice are derived according to Hogan, et al., "Manipulating the Mouse Embryo: A Laboratory Manual", Cold Spring Harbor Laboratory, which is incorporated herein by reference.

15   Embryonic stem cells are manipulated according to published procedures (Teratocarcinomas and embryonic stem cells: a practical approach, E.J. Robertson, ed., IRL Press, Washington, D.C., 1987; Zjilstra et al., Nature 342:435-438 (1989); and Schwartzberg et al., Science 246:799-803 (1989),

20 each of which is incorporated herein by reference).

       DNA cloning procedures are carried out according to J. Sambrook, et al. in Molecular Cloning: A Laboratory Manual, 2d ed., 1989, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., which is incorporated herein by

25 reference.

       Oligonucleotides are synthesized on an Applied Bio Systems oligonucleotide synthesizer according to specifications provided by the manufacturer.

       Hybridoma cells and antibodies are manipulated

30 according to "Antibodies: A Laboratory Manual", Ed Harlow and David Lane, Cold Spring Harbor Laboratory (1988), which is incorporated herein by reference.

EXAMPLE 1

Genomic Heavy Chain Human Ig Transgene

35   This Example describes the cloning and microinjection of a human genomic heavy chain immunoglobulin transgene which is microinjected into a murine zygote.

57

Nuclei are isolated from fresh human placental
tissue as described by Marzluff et al., "Transcription and
Translation: A Practical Approach", B.D. Hammes and
S.J. Higgins, eds., pp. 89-129, IRL Press, Oxford (1985)).
5 The isolated nuclei (or PBS washed human spermatocytes) are
embedded in a low melting point agarose matrix and lysed with
EDTA and proteinase κ to expose high molecular weight DNA,
which is then digested in the agarose with the restriction
enzyme NotI as described by M. Finney in Current Protocols in
10 Molecular Biology (F. Ausubel, et al., eds. John Wiley & Sons,
Supp. 4, 1988, Section 2.5.1).

The NotI digested DNA is then fractionated by pulsed
field gel electrophoresis as described by Anand et al.,
Nucl. Acids Res. 17:3425-3433 (1989). Fractions enriched for
15 the NotI fragment are assayed by Southern hybridization to
detect one or more of the sequences encoded by this fragment.
Such sequences include the heavy chain D segments, J segments,
μ and $\gamma1$ constant regions together with representatives of all
6 VH families (although this fragment is identified as 670 kb
20 fragment from HeLa cells by Berman et al. (1988), supra., we
have found it to be as 830 kb fragment from human placental an
sperm DNA). Those fractions containing this NotI fragment
(see Fig. 4) are pooled and cloned into the NotI site of the
vector pYACNN in Yeast cells. Plasmid pYACNN is prepared by
25 digestion of pYAC-4 Neo (Cook et al., Nucleic Acids Res. 16:
11817 (1988)) with EcoRI and ligation in the presence of the
oligonucleotide 5' - AAT TGC GGC CGC - 3'.

YAC clones containing the heavy chain NotI fragment
are isolated as described by Brownstein et al., Science
30 244:1348-1351 (1989), and Green et al., Proc. Natl. Acad. Sci.
USA 87:1213-1217 (1990), which are incorporated herein by
reference. The cloned NotI insert is isolated from high
molecular weight yeast DNA by pulse field gel electrophoresis
as described by M. Finney, op cit. The DNA is condensed by
35 the addition of 1 mM spermine and microinjected directly into
the nucleus of single cell embryos previously described.

58

EXAMPLE 2

Genomic κ Light Chain Human Ig Transgene Formed by In Vivo Homologous Recombination

A map of the human κ light chain has been described in Lorenz et al., Nucl. Acids Res. 15:9667-9677 (1987), which is incorporated herein by reference.

A 450 kb XhoI to NotI fragment that includes all of C_κ, the 3' enhancer, all J segments, and at least five different V segments is isolated and microinjected into the nucleus of single cell embryos as described in Example 1.

EXAMPLE 3

Genomic κ Light Chain Human Ig Transgene Formed by In Vivo Homologous Recombination

A 750 kb MluI to NotI fragment that includes all of the above plus at least 20 more V segments is isolated as described in Example 1 and digested with BssHII to produce a fragment of about 400 kb.

The 450 kb XhoI to NotI fragment plus the approximately 400 kb MluI to BssHII fragment have sequence overlap defined by the BssHII and XhoI restriction sites. Homologous recombination of these two fragments upon microinjection of a mouse zygote results in a transgene containing at least an additional 15-20 V segments over that found in the 450 kb XhoI/NotI fragment (Example 2).

EXAMPLE 4

Construction of Heavy Chain Mini-Locus

A. Construction of pGP1 and pGP2

pBR322 is digested with EcoRI and StyI and ligated with the following oligonucleotides to generate pGP1 which contains a 147 base pair insert containing the restriction sites shown in Fig. 8. The general overlapping of these oligos is also shown in Fig. 9.

59

The oligonucleotides are:

oligo-1 5' - CTT GAG CCC GCC TAA TGA GCG GGC TTT
              TTT TTG CAT ACT GCG GCC - 3'

5

oligo-2 5' - GCA ATG GCC TGG ATC CAT GGC GCG CTA
              GCA TCG ATA TCT AGA GCT CGA GCA -3'

oligo-3 5' - TGC AGA TCT GAA TTC CCG GGT ACC AAG
10   CTT ACG CGT ACT AGT GCG GCC GCT -3'

oligo-4 5' - AAT TAG CGG CCG CAC TAG TAC GCG TAA
              GCT TGG TAC CGG GAA TTC - 3'

15 oligo-5 5' - CAG ATC TGC ATG CTC GAG CTC TAG ATA
                  TCG ATG CTA GCG CGC CAT GGA TCC - 3'

oligo-6 5' - AGG CCA TTG CGG CCG CAG TAT GCA AAA
              AAA AGC CCG CTC ATT AGG CGG GCT - 3'

20
        This plasmid contains a large polylinker flanked by
rare cutting NotI sites for building large inserts that can be
isolated from vector sequences for microinjection. The
plasmid is based on pBR322 which is relatively low copy
25 compared to the pUC based plasmids (pGP1 retains the pBR322
copy number control region near the origin of replication).
Low copy number reduces the potential toxicity of insert
sequences. In addition, pGP1 contains a strong transcription
terminator sequence derived from trpA (Christie et al., Proc.
30 Natl. Acad. Sci. USA 78:4180 (1981)) inserted between the
ampicillin resistance gene and the polylinker. This further
reduces the toxicity associated with certain inserts by
preventing readthrough transcription coming from the
ampicillin promoters.

35   Plasmid pGP2 is derived from pGP1 to introduce an
additional restriction site (SfiI) in the polylinker. pGP1 is
digested with MluI and SpeI to cut the recognition sequences
in the polylinker portion of the plasmid.

        The following adapter oligonucleotides are ligated
40 to the thus digested pGP1 to form pGP2.

        5' CGC GTG GCC GCA ATG GCC A 3'
        5' CTA GTG GCC ATT GCG GCC A 3'

60

pGP2 is identical to pGP1 except that it contains an additional Sfi I site located between the MluI and SpeI sites. This allows inserts to be completely excised with SfiI as well as with NotI.

5

B.   Construction of pRE3 (rat enhancer 3')

An enhancer sequence located downstream of the rat constant region is included in the heavy chain constructs.

The heavy chain region 3' enhancer described by

10   Petterson et al., Nature 344:165-168 (1990), which is incorporated herein by reference) is isolated and cloned.   The rat IGH 3' enhancer sequence is PCR amplified by using the following oligonucleotides:

15   5' CAG GAT CCA GAT ATC AGT ACC TGA AAC AGG GCT TGC 3'
     5' GAG CAT GCA CAG GAC CTG GAG CAC ACA CAG CCT TCC 3'

The thus formed double stranded DNA encoding the 3' enhancer is cut with BamHI and SphI and clone into BamHI/SphI

20   cut pGP2 to yield pRE3 (rat enhancer 3').

C.   Cloning of Human J-μ Region

A substantial portion of this region is cloned by combining two or more fragments isolated from phage lambda

25   inserts.   See Fig. 9.

A 6.3 kb BamHI/HindIII fragment that includes all human J segments (Matsuda et al., EMBO J., 7:1047-1051 (1988); Ravetech et al.m Cell, 27:583-591 (1981), which are incorporated herein by reference) is isolated from human

30   genomic DNA library using the oligonucleotide GGA CTG TGT CCC TGT GTG ATG CTT TTG ATG TCT GGG GCC AAG.

An adjacent 10 kb HindIII/BamII fragment that contains enhancer, switch and constant region coding exons (Yasui et al., Eur. J. Immunol. 19:1399-1403 (1989)) is

35   similarly isolated using the oligonucleotide:
     CAC CAA GTT GAC CTG CCT GGT CAC AGA CCT GAC CAC CTA TGA

An adjacent 3' 1.5 kb BamHI fragment is similarly isolated using clone pMUM insert as probe (pMUM is 4 kb

61

EcoRI/HindIII fragment isolated from human genomic DNA library
with oligonucleotide:

       CCT GTG GAC CAC CGC CTC CAC CTT CAT

       CGT CCT CTT CCT CCT

5   mu membrane exon 1) and cloned into pUC19.

       pGP1 is digested with BamHI and BglII followed by
treatment with calf intestinal alkaline phosphatase.

       Fragments (a) and (b) from Fig. 9 are cloned in the
digested pGP1. A clone is then isolated which is oriented
10 such that 5' BamHI site is destroyed by BamHI/Bgl fusion. It
is identified as pMU (see Fig. 10). pMU is digested with
BamHI and fragment (c) from Fig. 9 is inserted. The
orientation is checked with HindIII digest. The resultant
plasmid pHIG1 (Fig. 10) contains an 18 kb insert encoding J
15 and Cμ segments.

D.   Cloning of Cμ Region

      pGP1 is digested with BamHI and HindIII is followed
by treatment with calf intestinal alkaline phosphatase (Fig.
20 14). The so treated fragment (b) of Fig. 14 and fragment (c)
of Fig. 14 are cloned into the BamHI/HindIII cut pGP1. Proper
orientation of fragment (c) is checked by HindIII digestion to
form pCON1 containing a 12 kb insert encoding the Cμ region.

      Whereas pHIG1 contains J segments, switch and μ
25 sequences in its 18 kb insert with an SfiI 3' site and a SpeI
5' site in a polylinker flanked by NotI sites, will be used
for rearranged VDJ segments. pCON1 is identical except that
it lacks the J region and contains only a 12 kb insert. The
use of pCON1 in the construction of fragment containing
30 rearranged VDJ segments will be described hereinafter.

E.   Cloning of γ-1 Constant Region (pREG2)

      The cloning of the human γ-1 region is depicted in
Fig. 16.

35   Yamamura et al., Proc. Natl. Acad. Sci. USA
83:2152-2156 (1986) reported the expression of membrane bound
human γ-1 from a transgene construct that had been partially
deleted on integration. Their results indicate that the 3'

62

BamHI site delineates a sequence that includes the
transmembrane rearranged and switched copy of the gamma gene
with a V-C intron of less than 5kb. Therefore, in the
unrearranged, unswitched gene, the entire switch region is
5 included in a sequence beginning less than 5 kb from the 5'
end of the first γ-1 constant exon. Therefore it is included
in the 5' 5.3 kb HindIII fragment (Ellison et al., Nucleic
Acids Res. 10:4071-4079 (1982), which is incorporated herein
by reference). Takahashi et al., Cell 29: 671-679 (1982),
10 which is incorporated herein by reference, also reports that
this fragment contains the switch sequence, and this fragment
together with the 7.7 kb HindIII to BamHI fragment must
include all of the sequences we need for the transgene
construct. An intronic sequence is a nucleotide sequence of
15 at least 15 contiguous nucleotides that occurs in an intron of
a specified gene.

Phage clones containing the γ-1 region are
identified and isolated using the following oligonucleotide
which is specific for the third exon of γ-I (CH3).

20

5' TGA GCC ACG AAG ACC CTG AGG
TCA AGT TCA ACT GGT ACG TGG 3'

A 7.7 kb HindIII to BglII fragment (fragment (a) in
25 Fig. 11) is cloned into HindIII/BglII cut pRE3 to form pREG1.
The upstream 5.3 kb HindIII fragment (fragment (b) in Fig. 11)
is cloned into HindIII digested pREG1 to form pREG2. Correct
orientation is confirmed by BamHI/SpeI digestion.

30 F. Combining Cγ and Cμ

·The previously described plasmid pHIG1 contains
human J segments and the Cμ constant region exons. To provide
a transgene containing the Cμ constant region gene segments,
pHIG1 was digested with SfiI (Fig. 10). The plasmid pREG2 was
35 also digested with SfiI to produce a 13.5 kb insert containing
human Cγ exons and the rat 3' enhancer sequence. These
sequences were combined to produce the plasmid pHIG3' (Fig.
12) containing the human J segments, the human Cμ constant

63

region, the human Cγ1 constant region and the rat 3' enhancer contained on a 31.5 kb insert.

A second plasmid encoding human Cμ and human Cγ1 without J segments is constructed by digesting pCON1 with SfiI and combining that with the SfiI fragment containing the human Cγ region and the rat 3' enhancer by digesting pREG2 with SfiI. The resultant plasmid, pCON (Fig. 12) contains a 26 kb NotI/SpeI insert containing human Cμ, human γ1 and the rat 3' enhancer sequence.

G. Cloning of D Segment

The strategy for cloning the human D segments is depicted in Fig. 13. Phage clones from the human genomic library containing D segments are identified and isolated using probes specific for diversity region sequences (Ichihara et al., EMBO J. 7:4141-4150 (1988)). The following oligonucleotides are used:

DXP1:   5' - TGG TAT TAC TAT GGT TCG GGG AGT TAT TAT
                    AAC CAC AGT GTC - 3'

DXP4:   5' - GCC TGA AAT GGA GCC TCA GGG CAC AGT GGG
                    CAC GGA CAC TGT - 3'

DN4:   5' - GCA GGG AGG ACA TGT TTA GGA TCT GAG GCC
                    GCA CCT GAC ACC - 3'

A 5.2 kb XhoI fragment (fragment (b) in Fig. 13) containing DLR1, DXP1, DXP'1, and DA1 is isolated from a phage clone identified with oligo DXP1.

A 3.2 kb XbaI fragment (fragment (c) in Fig. 13) containing DXP4, DA4 and DK4 is isolated from a phage clone identified with oligo DXP4.

Fragments (b), (c) and (d) from Fig. 13 are combined and cloned into the XbaI/XhoI site of pGP1 to form pHIG2 which contains a 10.6 kb insert.

This cloning is performed sequentially. First, the 5.2 kb fragment (b) in Fig. 13 and the 2.2 kb fragment (d) of

64

Fig. 13 are treated with calf intestinal alkaline phosphatase and cloned into pGP1 digested with XhoI and XbaI. The resultant clones are screened with the 5.2 and 2.2 kb insert. Half of those clones testing positive with the 5.2 and 2.2 kb
5 inserts have the 5.2 kb insert in the proper orientation as determined by BamHI digestion. The 3.2 kb XbaI fragment from Fig. 13 is then cloned into this intermediate plasmid containing fragments (b) and (d) to form pHIG2. This plasmid contains diversity segments cloned into the polylinker with a
10 unique 5' SfiI site and unique 3' SpeI site. The entire polylinker is flanked by NotI sites.

H.   Construction of Heavy Chain Minilocus

15   The following describes the construction of a human heavy chain mini-locus which contain one or more V segments.
            An unrearranged V segment corresponding to that identified as the V segment contained in the hybridoma of Newkirk et al., J. Clin. Invest. 81:1511-1518 (1988), which
20 is incorporated herein by reference, is isolated using the following oligonucleotide:

5' - GAT CCT GGT TTA GTT AAA GAG GAT TTT
              ATT CAC CCC TGT GTC - 3'

25
            A restriction map of the unrearranged V segment is determined to identify unique restriction sites which provide upon digestion a DNA fragment having a length approximately 2 kb containing the unrearranged V segment together with 5' and
30 3' flanking sequences. The 5' prime sequences will include promoter and other regulatory sequences whereas the 3' flanking sequence provides recombination sequences necessary for V-DJ joining. This approximately 3.0 kb V segment insert is cloned into the polylinker of pGB2 to form pVH1.
35   pVH1 is digested with SfiI and the resultant fragment is cloned into the SfiI site of pHIG2 to form a pHIG5'. Since pHIG2 contains D segments only, the resultant pHIG5' plasmid contains a single V segment together with D

65

segments. The size of the insert contained in pHIG5 is 10.6
kb plus the size of the V segment insert.

5   The insert from pHIG5 is excised by digestion with
NotI and SpeI and isolated. pHIG3' which contains J, Cμ and
cγ1 segments is digested with SpeI and NotI and the 3' kb
fragment containing such sequences and the rat 3' enhancer
sequence is isolated. These two fragments are combined and
ligated into NotI digested pGP1 to produce pHIG which contains
insert encoding a V segment, nine D segments, six functional J
10   segments, Cμ, Cγ and the rat 3' enhancer. The size of this
insert is approximately 43 kb plus the size of the V segment
insert.

I.   Construction of Heavy Chain Minilocus
15   by Homologous Recombination

As indicated in the previous section, the insert of
pHIG is approximately 43 to 45 kb when a single V segment is
employed. This insert size is at or near the limit of that
which may be readily cloned into plasmid vectors. In order to
20   provide for the use of a greater number of V segments, the
following describes in vivo homologous recombination of
overlapping DNA fragments which upon homologous recombination
within a zygote or ES cell form a transgene containing the rat
3' enhancer sequence, the human Cμ, the human Cγ1, human J
25   segments, human D segments and a multiplicity of human V
segments.

A 6.3 kb BamHI/HindIII fragment containing human J
segments (see fragment (a) in Fig. 9) is cloned into MluI/SpeI
digested pHIG5' using the following adapters:

30
                    5' GAT CCA AGC AGT 3'

                    5' CTA GAC TGC TTG 3'

                    5' CGC GTC GAA CTA 3'
35
                    5' AGC TTA GTT CGA 3'

The resultant is plasmid designated pHIG5'O
40   (overlap). The insert contained in this plasmid contains

66

human V, D and J segments. When the single V segment from
pVH1 is used, the size of this insert is approximately 17 kb
plus 2 kb. This insert is isolated and combined with the
insert from pHIG3' which contains the human J, Cμ, γ1 and rat
5   3' enhancer sequences. Both inserts contain human J segments
which provide for approximately 6.3 kb of overlap between the
two DNA fragments. When coinjected into the mouse zygote, in
vivo homologous recombination occurs generating a transgene
equivalent to the insert contained in pHIG.

10   This approach provides for the addition of a
multiplicity of V segments into the transgene formed in vivo.
For example, instead of incorporating a single V segment into
pHIG5', a multiplicity of V segments contained on (1) isolated
genomic DNA, (2) ligated DNA derived from genomic DNA, or (3)
15   DNA encoding a synthetic V segment repertoire is cloned into
pHIG2 at the SfiI site to generate pHIG5' V_N. The J segments
fragment (a) of Fig. 9 is then cloned into pHIG5' V_N and the
insert isolated. This insert now contains a multiplicity of V
segments and J segments which overlap with the J segments
20   contained on the insert isolated from pHIG3'. When
cointroduced into the nucleus of a mouse zygote, homologous
recombination occurs to generate in vivo the transgene
encoding multiple V segments and multiple J segments, multiple
D segments, the Cμ region, the Cγ1 region (all from human) and
25   the rat 3' enhancer sequence.

EXAMPLE 5
Construction of Light Chain Minilocus

A. Construction of pEμ1

30   The construction of pEμ1 is depicted in Fig. 16.
The mouse heavy chain enhancer is isolated on the XbaI to
EcoRI 678 bp fragment (Banerji et al., Cell 33:729-740 (1983))
from phage clones using oligo:

35   5' GAA TGG GAG TGA GGC TCT CTC ATA CCC
            TAT TCA GAA CTG ACT 3'

67

This Eμ fragment is cloned into EcoRV/XbaI digested pGP1 by blunt end filling in EcoRI site. The resultant plasmid is designated pEmu1.

B. Construction Of κ Light chain Minilocus

The κ construct contains at least one human V_κ segment, all five human J_κ segments, the human J-C_κ enhancer, human κ constant region exon, and, ideally, the human 3' κ enhancer (Meyer et al., EMBO J. 8:1959-1964 (1989)). The κ enhancer in mouse is 9 kb downstream from C_κ. However, it is as yet unidentified in the human. In addition, the construct contains a copy of the mouse heavy chain J-Cμ enhancers.

The minilocus is constructed from four component fragments:

(a) A 16 kb SmaI fragment that contains the human C_κ exon and the 3' human enhancer by analogy with the mouse locus;

(b) A 5' adjacent 5 kb SmaI fragment, which contains all five J segments;

(c) The mouse heavy chain intronic enhancer isolated from pEμ1 (this sequence is included to induce expression of the light chain construct as early as possible in B-cell development. Because the heavy chain genes are transcribed earlier than the light chain genes, this heavy chain enhancer is presumably active at an earlier stage than the intronic κ enhancer); and

(d) A fragment containing one or more V segments.

The preparation of this construct is as follows. Human placental DNA is digested with SmaI and fractionated on agarose gel by electrophoresis. Similarly, human placental DNA is digested with BamHI and fractionated by electrophoresis. The 16 kb fraction is isolated from the SmaI digested gel and the 11 kb region is similarly isolated from the gel containing DNA digested with BamHI.

The 16 kb SmaI fraction is cloned into Lambda FIX II (Stratagene, La Jolla, California) which has been digested with XhoI, treated with klenow fragment DNA polymerase to fill in the XhoI restriction digest product. Ligation of the 16 kb

68

SmaI fraction destroys the SmaI sites and lases XhoI sites
intact.

The 11 kb BamHI fraction is cloned into λ EMBL3
(Strategene, La Jolla, California) which is digested with
5 BamHI prior to cloning.

Clones from each library were probed with the Cκ
specific oligo:

5' GAA CTG TGG CTG CAC CAT CTG TCT
    TCA TCT TCC CGC CAT CTG 3'

10

A 16 kb XhoI insert that was subcloned into the XhoI
cut pEμ1 so that Cκ is adjacent to the SmaI site. The
15 resultant plasmid was designated pKap1.

The above Cκ specific oligonucleotide is used to
probe the λ EMBL3/BamHI library to identify an 11 kb clone. A
5 kb SmaI fragment (fragment (b) in Fig. 20) is subcloned and
subsequently inserted into pKap1 digested with SmaI. Those
20 plasmids containing the correct orientation of J segments, Cκ
and the Eμ enhancer are designated pKap2.

One or more Vκ segments are thereafter subcloned
into the MluI site of pKap2 to yield the plasmid pKapH which
encodes the human Vκ segments, the human Jκ segments, the
25 human Cκ segments and the human Eμ enhancer. This insert is
excised by digesting pKapH with NotI and purified by agarose
gel electrophoresis. The thus purified insert is
microinjected into the pronucleus of a mouse zygote as
previously described.

30
C.   Construction of κ Light Chain Minilocus by
      In Vivo Homologous Recombination

The 11 kb BamHI fragment is cloned into BamHI
35 digested pGP1 such that the 3' end is toward the SfiI site.
The resultant plasmid is designated pKAPint. One or more Vκ
segments is inserted into the polylinker between the BamHI and
SpeI sites in pKAPint to form pKapHV. The insert of pKapHV is
excised by digestion with NotI and purified. The insert from
40 pKap2 is excised by digestion with NotI and purified. Each of
these fragments contain regions of homology in that the

69

fragment from pKapHV contains a 5 kb sequence of DNA that
include the J_k segments which is substantially homologous to
the 5 kb SmaI fragment contained in the insert obtained from
pKap2. As such, these inserts are capable of homologously

5   recombining when microinjected into a mouse zygote to form a
transgene encoding V_k, J_k and C_k.

EXAMPLE 6

10   Isolation of Genomic Clones
                Corresponding to Rearranged and Expressed
                Copies of Immunoglobulin k Light Chain Genes

          This example describes the cloning of immunoglobulin

15 k light chain genes from cultured cells that express an
immunoglobulin of interest. Such cells may contain multiple
alleles of a given immunoglobulin gene. For example, a
hybridoma might contain four copies of the k light chain gene,
two copies from the fusion partner cell line and two copies

20 from the original B-cell expressing the immunoglobulin of
interest. Of these four copies, only one encodes the
immunoglobulin of interest, despite the fact that several of
them may be rearranged. The procedure described in this
example allows for the selective cloning of the expressed copy

25 of the k light chain.

A.   Double Stranded cDNA
          Cells from human hybridoma, or lymphoma, or other
cell line that synthesizes either cell surface or secreted or

30 both forms of IgM with a k light chain are used for the
isolation of polyA+ RNA. The RNA is then used for the
synthesis of oligo dT primed cDNA using the enzyme reverse
transcriptase. The single stranded cDNA is then isolated and
G residues are added to the 3' end using the enzyme

35 polynucleotide terminal transferase. The Gtailed
single-stranded cDNA is then purified and used as template for
second strand synthesis (catalyzed by the enzyme DNA
polymerase) using the following oligonucleotide as a primer:

40   5' - GAG GTA CAC TGA CAT ACT GGC ATG CCC
                     CCC CCC CCC - 3'

70

The double stranded cDNA is isolated and used for determining the nucleotide sequence of the 5' end of the mRNAs encoding the heavy and light chains of the expressed immunoglobulin molecule. Genomic clones of these expressed

5 genes are then isolated. The procedure for cloning the expressed light chain gene is outlined in part B below.

B. Light Chain

The double stranded cDNA described in part A is

10 denatured and used as a template for a third round of DNA synthesis using the following oligonucleotide primer:

      5' - GTA CGC CAT ATC AGC TGG ATG AAG TCA TCA GAT
            GGC GGG AAG ATG AAG ACA GAT GGT GCA - 3'

15

This primer contains sequences specific for the constant portion of the κ light chain message (TCA TCA GAT GGC GGG AAG ATG AAG ACA GAT GGT GCA) as well as unique sequences

20 that can be used as a primer for the PCR amplification of the newly synthesized DNA strand (GTA CGC CAT ATC AGC TGG ATG AAG). The sequence is amplified by PCR using the following two oligonucleotide primers:

                  5' - GAG GTA CAC TGA CAT ACT GGC ATG -3'
25   5' - GTA CGC CAT ATC AGC TGG ATG AAG -3'

The PCR amplified sequence is then purified by gel

30 electrophoresis and used as template for dideoxy sequencing reactions using the following oligonucleotide as a primer:

          5' - GAG GTA CAC TGA CAT ACT GGC ATG -3'

The first 42 nucleotides of sequence will then be

35 used to synthesize a unique probe for isolating the gene from which immunoglobulin message was transcribed. This synthetic 42 nucleotide segment of DNA will be referred to below as o-kappa.

40   A Southern blot of DNA, isolated from the Ig expressing cell line and digested individually and in pairwise combinations with several different restriction endonucleases

71

including SmaI, is then probed with the 32-P labelled unique oligonucleotide o-kappa. A unique restriction endonuclease site is identified upstream of the rearranged V segment.

5   DNA from the Ig expressing cell line is then cut with SmaI and second enzyme (or BamHI or KpnI if there is SmaI site inside V segment). Any resulting non-blunted ends are treated with the enzyme T4 DNA polymerase to give blunt ended DNA molecules. Then add restriction site encoding linkers (BamHI, EcoRI or XhoI depending on what site does not exist in

10   fragment) and cut with the corresponding linker enzyme to give DNA fragments with BamHI, EcoRI or XhoI ends. The DNA is then size fractionated by agarose gel electrophoresis, and the fraction including the DNA fragment covering the expressed V segment is cloned into lambda EMBL3 or Lambda FIX (Stratagene,

15   La Jolla, California). V segment containing clones are isolated using the unique probe o-kappa. DNA is isolated from positive clones and subcloned into the polylinker of pKap1. The resulting clone is called pRKL.

20   EXAMPLE 7
                       Isolation of Genomic Clones
                 Corresponding to Rearranged Expressed Copies
                 of Immunoglobulin Heavy Chain μ Genes

       This example describes the cloning of immunoglobulin

25   heavy chain μ genes from cultured cells of expressed and immunoglobulin of interest. The procedure described in this example allows for the selective cloning of the expressed copy of a μ heavy chain gene.

       Double-stranded cDNA is prepared and isolated as

30   described herein before. The double-stranded cDNA is denatured and used as a template for a third round of DNA synthesis using the following oligonucleotide primer:

           5' - GTA CGC CAT ATC AGC TGG ATG AAG ACA GGA GAC

35   GAG GGG GAA AAG GGT TGG GGC GGA TGC - 3'

       This primer contains sequences specific for the constant portion of the μ heavy chain message (ACA GGA GAC GAG GGG GAA AAG GGT TGG GGC GGA TGC) as well as unique sequences

72

that can be used as a primer for the PCR amplification of the
newly synthesized DNA strand (GTA CGC CAT ATC AGC TGG ATG
AAG). The sequence is amplified by PCR using the following
two oligonucleotide primers:

5

          5' – GAG GTA CAC TGA CAT ACT GGC ATG – 3'
          5' – GTA CTC CAT ATC AGC TGG ATG AAG – 3'

      The PCR amplified sequence is then purified by gel
10 electrophoresis and used as template for dideoxy sequencing
reactions using the following oligonucleotide as a primer:

      5' – GAG GTA CAC TGA CAT ACT GGC ATG – 3'

15   The first 42 nucleotides of sequence are then used
to synthesize a unique probe for isolating the gene from
which immunoglobulin message was transcribed. This synthetic
42 nucleotide segment of DNA will be referred to below as
o-mu.

20   A Southern blot of DNA, isolated from the Ig
expressing cell line and digested individually and in pairwise
combinations with several different restriction endonucleases
including MluI (MluI is a rare cutting enzyme that cleaves
between the J segment and mu CH1), is then probed with the
25 32-P labelled unique oligonucleotide o-mu. A unique
restriction endonuclease site is identified upstream of the
rearranged V segment.

      DNA from the Ig expressing cell line is then cut
with MluI and second enzyme. MluI or SpeI adapter linkers are
30 then ligated onto the ends and cut to convert the upstream
site to MluI or SpeI. The DNA is then size fractionated by
agarose gel electrophoresis, and the fraction including the
DNA fragment covering the expressed V segment is cloned
directly into the plasmid pGPI. V segment containing clones
35 are isolated using the unique probe o-mu, and the insert is
subcloned into MluI or MluI/SpeI cut plasmid pCON2. The
resulting plasmid is called pRMGH.

73

EXAMPLE 8

Construction of Human κ Miniloci Transgenes

Light Chain Minilocus

A human genomic DNA phage library was screened with

5   kappa light chain specific oligonucleotide probes and isolated
clones spanning the J_K-C region. A 5.7 kb ClaI/XhoI fragment
containing J_K1 together with a 13 kb XhoI fragment containing
J_K2-5 and C_K into pGP1d was cloned and used to create the
plasmid pKcor. This plasmid contains J_K1-5, the kappa

10   intronic enhancer and C_K together with 4.5 kb of 5' and 9 kb
of 3' flanking sequences. It also has a unique 5' XhoI site
for cloning V_K segments and a unique 3' SalI site for
inserting additional cis-acting regulatory sequences.

15   V kappa genes

A human genomic DNA phage library was screened with
V_K light chain specific oligonucleotide probes and isolated
clones containing human V_K segments. Functional V segments
were identified by DNA sequence analysis. These clones

20   contain TATA boxes, open reading frames encoding leader and
variable peptides (including 2 cysteine residues), splice
sequences, and recombination heptamer-12 bp spacer-nonamer
sequences. Three of the clones were mapped and sequenced.
Two of the clones, 65.5 and 65.8 appear to be functional, they

25   contain TATA boxes, open reading frames encoding leader and
variable peptides (including 2 cysteine residues), splice
sequences, and recombination heptamer-12 bp spacer-nonamer
sequences. The third clone, 65.4, appears to encode a V_KI
pseudogene as it contains a non-canonical recombination

30   heptamer.

One of the functional clones, Vk 65-8, which encodes
a VkIII family gene, was used to build a light chain minilocus
construct.

35   pKC1

The kappa light chain minilocus transgene pKC1 (Fig.
32) was generated by inserting a 7.5 kb XhoI/SalI fragment

74

containing V$_K$ 65.8 into the 5' XhoI site of pKcor. The transgene insert was isolated by digestion with NotI prior to injection.

5   The purified insert was microinjected into the pronuclei of fertilized (C57BL/6 x CBA)F2 mouse embryos and transferred the surviving embryos into pseudopregnant females as described by Hogan et al. (in Methods of Manipulating the Mouse Embryo, 1986, Cold Spring Harbor Laboratory, New York). Mice that developed from injected embryos were analyzed for

10   the presence of transgene sequences by Southern blot analysis of tail DNA. Transgene copy number was estimated by band intensity relative to control standards containing known quantities of cloned DNA. Serum was isolated from these animals and assayed for the presence of transgene encoded

15   human Ig kappa protein by ELISA as described by Harlow and Lane (in Antibodies: A Laboratory Manual, 1988, Cold Spring Harbor Laboratory, New York). Microtiter plate wells were coated with mouse monoclonal antibodies specific for human Ig kappa (clone 6E1, #0173, AMAC, Inc., Westbrook, ME), human IgM

20   (Clone AF6, #0285, AMAC, Inc., Westbrook, ME) and human IgG1 (clone JL512, #0280, AMAC, Inc., Westbrook, ME). Serum samples were serially diluted into the wells and the presence of specific immunoglobulins detected with affinity isolated alkaline phosphatase conjugated goat anti-human Ig

25   (polyvalent) that had been pre-adsorbed to minimize cross-reactivity with mouse immunoglobulins.

         Fig. 35 shows the results of an ELISA assay of serum from 8 mice (I.D. #676, 674, 673, 670, 666, 665, 664, and 496). The first seven of these mice developed from embryos

30   that were injected with the pKC1 transgene insert and the eighth mouse is derived from a mouse generated by microinjection of the pHC1 transgene (described previously). Two of the seven mice from KC1 injected embryos (I.D.#'s 666 and 664) did not contain the transgene insert as assayed by

35   DAN Southern blot analysis, and five of the mice (I.D.#'s 676, 674, 673, 670, and 665) contained the transgene. All but one of the KC1 transgene positive animals express detectable levels of human Ig kappa protein, and the single non-

expressing animal appears to be a genetic mosaic on the basis
of DNA Southern blot analysis. The pHC1 positive transgenic
mouse expresses human IgM and IgG1 but not Ig kappa,
demonstrating the specificity of the reagents used in the
5   assay.

pKC2

The kappa light chain minilocus transgene pKC2 was
generated by inserting an 8 kb XhoI/SalI fragment containing
10 V_κ 65.5 into the 5' XhoI site of pKC1. The resulting
transgene insert, which contains two V_κ segments, was isolated
prior to microinjection by digestion with NotI.

15

pKVe2

. This construct is identical to pKC1 except that it
includes 1.2 kb of additional sequence 5' of J_κ and is missing
4.5 kb of sequence 3' of V_κ 65.8. In additional it contains a
20 0.9 kb XbaI fragment containing the mouse heavy chain J-m
intronic enhancer (Banerji et al., Cell 33:729-740 (1983))
together with a 1.4 kb MluI HindIII fragment containing the
human heavy chain J-m intronic enhancer (Hayday et al., Nature
307:334-340 (1984)) inserted downstream. This construct tests
25 the feasibility of initiating early rearrangement of the light
chain minilocus to effect allelic and isotypic exclusion.
Analogous constructs can be generated with different
enhancers, i.e., the mouse or rat 3' kappa or heavy chain
enhancer (Meyer and Neuberger, EMBO J. 8:1959-1964 (1989);
30 Petterson et al. Nature 344:165-168 (1990), which are
incorporated herein by reference).

Rearranged Light Chain Transgenes

A kappa light chain expression cassette was designed
35 to reconstruct functionally rearranged light chain genes that
have been amplified by PCR from human B-cell DNA. The scheme
is outlined in Fig. 33. PCR amplified light chain genes are
cloned into the vector pK5nx that includes 3.7 kb of 5'

76

flanking sequences isolated from the kappa light chain gene
65.5. The VJ segment fused to the 5' transcriptional
sequences are then cloned into the unique XhoI site of the
vector pK31s that includes $J_\kappa 2-4$, the J_κ intronic enhancer, C_κ,

5   and 9 kb of downstream sequences. The resulting plasmid
contains a reconstructed functionally rearranged kappa light
chain transgene that can be excised with NotI for
microinjection into embryos. The plasmids also contain unique
SalI sites at the 3' end for the insertion of additional cis-

10 acting regulatory sequences.

        Two synthetic oligonucleotides (o-130, o-131) were
used ɔ amplify rearranged kappa light chain genes from human
spleen genomic DNA. Oligonucleotide o-131 (gga ccc aga
(g,c)gg aac cat gga a(g,a)(g,a,t,c)) is complementary to the

15 5' region of $V_\kappa III$ family light chain genes and overlaps the
first ATC of the leader sequence. Oligonucleotide o-130 (gtg
caa tca att ctc gag ttt gac tac aga c) is complementary to a
sequence appr.ximately 150 bp 3' of $J_\kappa 1$ and includes an XhoI
site. These two oligonucleotides amplify a 0.7 kb DNA

20 fragment from human spleen DNA corresponding to rearranged
$V_\kappa III$ genes joined to $J_\kappa 1$ segments. The PCR amplified DNA was
digested with NcoI and XhoI and cloned individual PCR products
into the plasmid pNN03. The DNA sequence of 5 clones was
determined and identified two with functional VJ joints (open

25 reading frames). Additional functionally rearranged light
chain clones are collected. The functionally rearranged clones
can be individually cloned into light chain expression
cassette described above (Fig. 33). Transgenic mice generated
with the rearranged light chain constructs can be bred with

30 heavy chain minilocus transgenics to produce a strain of mice
that express a spectrum of fully human antibodies in which all
of the diversity of the primary repertoire is contributed by
the heavy chain. One source of light chain diversity can be
from somatic mutation. Because not all light chains will be

35 equivalent with respect to their ability to combine with a
variety of different heavy chains, different strains of mice,
each containing different light chain constructs can be
generated and tested. The advantage of this scheme, as

77

opposed to the use of unrearranged light chain miniloci, is
the increased light chain allelic and isotypic exclusion that
comes from having the light chain ready to pair with a heavy
chain as soon as heavy chain VDJ joining occurs.   This

5 combination can result in an increased frequency of B-cells
expressing fully human antibodies, and thus it can facilitate
the isolation of human Ig expressing hybridomas.

        NotI inserts of plasmids pIGM1, pHC1, pIGG1, pKC1,
and pKC2 were isolated away from vector sequences by agarose

10 gel electrophoresis.   The purified inserts were microinjected
into the pronuclei of fertilized (C57BL/6 x CBA)F2 mouse
embryos and transferred the surviving embryos into
pseudopregnant females as described by Hogan et al. (Hogan et
al., Methods of Manipulating the Mouse Embryo, Cold Spring

15 Harbor Laboratory, New York (1986)).

                            EXAMPLE 9
Inactivation of the Mouse Kappa Light Chain Gene by Homologous
Recombination

20
        This example describes the inactivation of the mouse
endogenous kappa locus by homologous recombination in
embryonic stem (ES) cells followed by introduction of the
mutated gene into the mouse germ line by injection of targeted

25 ES cells bearing an inactivated kappa allele into early mouse
embryos (blastocysts).

        The strategy is to delete J_K and C_K by homologous
recombination with a vector containing DNA sequences
homologous to the mouse kappa locus in which a 4.5 kb segment

30 of the locus, spanning the J_K gene and C_K segments, is deleted
and replaced by the selectable marker neo.

Construction of the kappa targeting vector

        The plasmid pGEM7 (KJ1) contains the neomycin

35 resistance gene (neo), used for drug selection of transfected
ES cells, under the transcriptional control of the mouse
phosphoglycerate kinase (pgk) promoter (XbaI/TaqI fragment;
Adra et al., Gene 60:65-74 (1987)) in the cloning vector pGEM-
7Zf(+).   The plasmid also includes a heterologous

78

polyadenylation site for the neo gene, derived from the 3'
region of the mouse pgk gene (PvuII/HindIII fragment; Boer et
al., Biochemical Genetics, 28:299-308 (1990)). This plasmid
was used as the starting point for construction of the kappa
5   targeting vector. The first step was to insert sequences
homologous to the kappa locus 3' of the neo expression
cassette.

          Mouse kappa chain sequences (Fig. 20a) were isolated
from a genomic phage library derived from liver DNA using
10 oligonucleotide probes specific for the $C\kappa$ locus:

5'- GGC TGA TGC TGC ACC AAC TGT ATC CAT CTT CCC ACC ATC CAG
-3'
and for the $J\kappa5$ gene segment:
15 5'- CTC ACG TTC GGT GCT GGG ACC AAG CTG GAG CTG AAA CGT AAG -
3'.

          An 8 kb BglII/SacI fragment extending 3' of the
mouse C_K segment was isolated from a positive phage clone in
20 two pieces, as a 1.2 kb BglII/SacI fragment and a 6.8 kb SacI
fragment, and subcloned into BglII/SacI digested pGEM7 (KJ1)
to generate the plasmid pNEO-K3' (Fig. 20b).

          A 1.2 kb EcoRI/SphI fragment extending 5' of the J_K
region was also isolated from a positive phage clone. An
25 SphI/XbaI/BglII/EcoRI adaptor was ligated to the SphI site of
this fragment, and the resulting EcoRI fragment was ligated
into EcoRI digested pNEO-K3', in the same 5' to 3' orientation
as the neo gene and the downstream 3' kappa sequences, to
generate pNEO-K5'3' (Fig. 20c).

30   The Herpes Simplex Virus (HSV) thymidine kinase (TK)
gene was then included in the construct in order to allow for
enrichment of ES clones bearing homologous recombinants, as
described by Mansour et al., Nature 336:348-352 (1988), which
is incorporated herein by reference. The HSV TK cassette was
35 obtained from the plasmid pGEM7 (TK), which contains the
structural sequences for the HSV TK gene bracketed by the
mouse pgk promoter and polyadenylation sequences as described
above for pGEM7 (KJ1). The EcoRI site of pGEM7 (TK) was

79

modified to a BamHI site and the TK cassette was then excised
as a BamHI/HindIII fragment and subcloned into pGP1b to
generate pGP1b-TK. This plasmid was linearized at the XhoI
site and the XhoI fragment from pNEO-K5'3', containing the neo
5   gene flanked by genomic sequences from 5' of $J\kappa$ and 3' of $C\kappa$,
was inserted into pGP1b-TK to generate the targeting vector
J/C KI (Fig. 20d). The putative structure of the genomic
kappa locus following homologous recombination with J/C K1 is
shown in Fig. 20e.

10

Generation and analysis of ES cells with targeted inactivation
of a kappa allele

        The ES cells used were the AB-1 line grown on
mitotically inactive SNL76/7 cell feeder layers (McMahon and
15 Bradley, Cell 62:1073-1085 (1990)) essentially as described
(Robertson, E.J. (1987) in Teratocarcinomas and Embryonic Stem
Cells: A Practical Approach. E.J. Robertson, ed. (Oxford: IRL
Press), p. 71-112). Other suitable ES lines include, but are
not limited to, the E14 line (Hooper et al. (1987) Nature 326:
20 292-295), the D3 line (Doetschman et al. (1985) J. Embryol.
Exp. Morph. 87: 27-45), and the CCE line (Robertson et al.
(1986) Nature 323: 445-448). The success of generating a
mouse line from ES cells bearing a specific targeted mutation
depends on the pluripotence of the ES cells (i.e., their
25 ability, once injected into a host blastocyst, to participate
in embryogenesis and contribute to the germ cells of the
resulting animal).
        The pluripotence of any given ES cell line can vary
with time in culture and the care with which it has been
30 handled. The only definitive assay for pluripotence is to
determine whether the specific population of ES cells to be
used for targeting can give rise to chimeras capable of
germline transmission of the ES genome. For this reason,
prior to gene targeting, a portion of the parental population
35 of AB-1 cells is injected into C57B1/6J blastocysts to
ascertain whether the cells are capable of generating chimeric
mice with extensive ES cell contribution and whether the

80

majority of these chimeras can transmit the ES genome to
progeny.

5   The kappa chain inactivation vector J/C K1 was
digested with NotI and electroporated into AB-1 cells by the
methods described (Hasty et al., Nature, 350:243-246 (1991)).
Electroporated cells were plated onto 100 mm dishes at a
density of 1-2 x 10⁶ cells/dish. After 24 hours, G418
(200μg/ml of active component) and FIAU (0.5μM) were added to
the medium, and drug-resistant clones were allowed to develop
10   over 10-11 days. Clones were picked, trypsinized, divided
into two portions, and further expanded. Half of the cells
derived from each clone were then frozen and the other half
analyzed for homologous recombination between vector and
target sequences.

15   DNA analysis was carried out by Southern blot
hybridization. DNA was isolated from the clones as described
(Laird et al., Nucl. Acids Res. 19:4293 (1991)) digested with
XbaI and probed with the 800 bp EcoRI/XbaI fragment indicated
in Fig. 20e as probe A. This probe detects a 3.7 kb XbaI
20   fragment in the wild type locus, and a diagnostic 1.8 kb band
in a locus which has homologously recombined with the
targeting vector (see Fig. 20a and e). Of 901 G418 and FIAU
resistant clones screened by Southern blot analysis, 7
displayed the 1.8 kb XbaI band indicative of a homologous
25   recombination into one of the kappa genes. These 7 clones
were further digested with the enzymes BglII, SacI, and PstI
to verify that the vector integrated homologously into one of
the kappa genes. When probed with the diagnostic 800 bp
EcoRI/XbaI fragment (probe A), BglII, SacI, and PstI digests
30   of wild type DNA produce fragments of 4.1, 5.4, and 7 kb,
respectively, whereas the presence of a targeted kappa allele
would be indicated by fragments of 2.4, ⁻ 5, and 5.7 kb,
respectively (see Fig. 20a and e). All positive clones
detected by the XbaI digest showed the expected BglII, SacI,
35   and PstI restriction fragments diagnostic of a homologous
recombination at the kappa light chain. In addition, Southern
blot analysis of an NsiI digest of the targeted clones using a
neo specific probe (probe B, Fig. 20e) generated only the

81

predicted fragment of 4.2 kb, demonstrating that the clones each contained only a single copy of the targeting vector.

Generation of mice bearing the inactivated kappa chain

5   Five of the targeted ES clones described in the previous section were thawed and injected into C57B1/6J blastocysts as described (Bradley, A. (1987) in Teratocarcinomas and Embryonic Stem Cells: A Practical Approach. E.J. Robertson, ed. (Oxford: IRL Press), p. 113-151)

10   and transferred into the uteri of pseudopregnant females to generate chimeric mice resulting from a mixture of cells derived from the input ES cells and the host blastocyst. The extent of ES cell contribution to the chimeras can be visually estimated by the amount of agouti coat coloration, derived

15   from the ES cell line, on the black C57B1/6J background. Approximately half of the offspring resulting from blastocyst injection of the targeted clones were chimeric (i.e., showed agouti as well as black pigmentation) and of these, the majority showed extensive (70 percent or greater) ES cell

20   contribution to coat pigmentation. The AB1 ES cells are an XY cell line and a majority of these high percentage chimeras were male due to sex conversion of female embryos colonized by male ES cells. Male chimeras derived from 4 of the 5 targeted clones were bred with C57BL/6J females and the offspring

25   monitored for the presence of the dominant agouti coat color indicative of germline transmission of the ES genome. Chimeras from two of these clones consistently generated agouti offspring. Since only one copy of the kappa locus was targeted in the injected ES clones, each agouti pup had a 50

30   percent chance of inheriting the mutated locus. Screening for the targeted gene was carried out by Southern blot analysis of Bgl II-digested DNA from tail biopsies, using the probe utilized in identifying targeted ES clones (probe A, Fig. 20e). As expected, approximately 50 percent of the agouti

35   offspring showed a hybridizing Bgl II band of 2.4 kb in addition to the wild-type band of 4.1 kb, demonstrating the germline transmission of the targeted kappa locus.

82

In order to generate mice homozygous for the
mutation, heterozygotes were bred together and the kappa
genotype of the offspring determined as described above. As
expected, three genotypes were derived from the heterozygote
5 matings: wild-type mice bearing two copies of a normal kappa
locus, heterozygotes carrying one targeted copy of the kappa
gene and one NT kappa gene, and mice homozygous for the kappa
mutation. The deletion of kappa sequences from these latter
mice was verified by hybridization of the Southern blots with
10 a probe specific for J_K (probe C, Fig. 20a). Whereas
hybridization of the J_K probe was observed to DNA samples from
heterozygous and wild-type siblings, no hybridizing signal was
present in the homozygotes, attesting to the generation of a
novel mouse strain in which both copies of the kappa locus
15 have been inactivated by deletion as a result of targeted
mutation.

EXAMPLE 10

Inactivation of the Mouse Heavy Chain Gene by Homologous
20 Recombination

This example describes the inactivation of the
endogenous murine immunoglobulin heavy chain locus by
homologous recombination in embryonic stem (ES) cells. The
strategy is to delete the endogenous heavy chain J segments by
25 homologous recombination with a vector containing heavy chain
sequences from which the J_H region has been deleted and
replaced by the gene for the selectable marker neo.

Construction of a heavy chain targeting vector

30 Mouse heavy chain sequences containing the J_H region
(Fig. 21a) were isolated from a genomic phage library derived
from the D3 ES cell line (Gossler et al., Proc. Natl. Acad.
Sci. U.S.A. 83:9065-9069 (1986)) using a J_H4 specific
oligonucleotide probe:

35

5'- ACT ATG CTA TGG ACT ACT GGG GTC AAG GAA CCT CAG TCA CCG
-3'

83

A 3.5 kb genomic SacI/StuI fragment, spanning the J_H
region, was isolated from a positive phage clone and subcloned
into SacI/SmaI digested pUC18. The resulting plasmid was
designated pUC18 J_H. The neomycin resistance gene (neo), used
5   for drug selection of transfected ES cells, was derived from a
repaired version of the plasmid pGEM7 (KJ1). A report in the
literature (Yenofsky et al. (1990) <u>Proc. Natl. Acad. Sci.
(U.S.A.)</u> <u>87</u>: 3435-3439) documents a point mutation the neo
coding sequences of several commonly used expression vectors,
10 including the construct pMClneo (Thomas and Cappechi (1987)
<u>Cell</u> <u>51</u>: 503-512) which served as the source of the neo gene
used in pGEM7 (KJ1). This mutation reduces the activity of
the neo gene product and was repaired by replacing a
restriction fragment encompassing the mutation with the
15 corresponding sequence from a wild-type neo clone. The
HindIII site in the prepared pGEM7 (KJ1) was converted to a
SalI site by addition of a synthetic adaptor, and the neo
expression cassette excised by digestion with XbaI/SalI. The
ends of the neo fragment were then blunted by treatment with
20 the Klenow form of DNA polI, and the neo fragment was
subcloned into the NaeI site of pUC18 J_H, generating the
plasmid pUC18 J_H-neo (Fig. 21b).

Further construction of the targeting vector was
carried out in a derivative of the plasmid pGP1b. pGP1b was
25 digested with the restriction enzyme NotI and ligated with the
following oligonucleotide as an adaptor:

5'- GGC CGC TCG ACG ATA GCC TCG AGG CTA TAA ATC TAG AAG AAT
TCC AGC AAA GCT TTG GC -3'

30
The resulting plasmid, called pGMT, was used to
build the mouse immunoglobulin heavy chain targeting
construct.

The Herpes Simplex Virus (HSV) thymidine kinase (TK)
35 gene was included in the construct in order to allow for
enrichment of ES clones bearing homologous recombinants, as
described by Mansour et al. (<u>Nature</u> <u>336</u>, 348-352 (1988)). The
HSV TK gene was obtained from the plasmid pGEM7 (TK) by

84

digestion with EcoRI and HindIII. The TK DNA fragment was subcloned between the EcoRI and HindIII sites of pGMT, creating the plasmid pGMT-TK (Fig. 21c).

5   To provide an extensive region of homology to the target sequence, a 5.9 kb genomic XbaI/XhoI fragment, situated 5' of the J_H region, was derived from a positive genomic phage clone by limit digestion of the DNA with XhoI, and partial digestion with XbaI. As noted in Fig. 21a, this XbaI site is not present in genomic DNA, but is rather derived from phage

10 sequences immediately flanking the cloned genomic heavy chain insert in the positive phage clone. The fragment was subcloned into XbaI/XhoI digested pGMT-TK, to generate the plasmid pGMT-TK-J_H5' (Fig. 21d).

      The final step in the construction involved the

15 excision from pUC18 J_H-neo of the 2.8 kb EcoRI fragment which contained the neo gene and flanking genomic sequences 3' of J_H. This fragment was blunted by Klenow polymerase and subcloned into the similarly blunted XhoI site of pGMT-TK-J_H5'. The resulting construct, J_HKO1 (Fig. 21e),

20 contains 6.9 kb of genomic sequences flanking the J_H locus, with a 2.3 kb deletion spanning the J_H region into which has been inserted the neo gene. Fig. 21f shows the structure of an endogenous heavy chain gene after homologous recombination with the targeting construct.

25

EXAMPLE 11

Generation and analysis of targeted ES cells

      AB-1 ES cells (McMahon and Bradley, <u>Cell</u> <u>62</u>:1073-1085 (1990)) were grown on mitotically inactive

30 SNL76/7 cell feeder layers essentially as described (Robertson, E.J. (1987) <u>Teratocarcinomas and Embryonic Stem Cells: A Practical Approach</u>. E.J. Robertson, ed. (Oxford: IRL Press), pp. 71-112). As described in the previous example, prior to electroporation of ES cells with the targeting

35 construct J_HKO1, the pluripotency of the ES cells was determined by generation of AB-1 derived chimeras which were shown capable of germline transmission of the ES genome.

85

The heavy chain inactivation vector J_HKO1 was
digested with NotI and electroporated into AB-1 cells by the
methods described (Hasty et al., Nature 350:243-246 (1991)).
Electroporated cells were plated into 100 mm dishes at a
5   density of 1-2 x 10^6 cells/dish. After 24 hours, G418
(200mg/ml of active component) and FIAU (0.5mM) were added to
the medium, and drug-resistant clones were allowed to develop
over 8-10 days. Clones were picked, trypsinized, divided into
two portions, and further expanded. Half of the cells derived
10 from each clone were then frozen and the other half analyzed
for homologous recombination between vector and target
sequences.
        DNA analysis was carried out by Southern blot
hybridization. DNA was isolated from the clones as described
15 (Laird et al. (1991) Nucleic Acids Res. 19: 4293), digested
with StuI and probed with the 500 bp EcoRI/StuI fragment
designated as probe A in Fig. 21f. This probe detects a StuI
fragment of 4.7 kb in the wild-type locus, whereas a 3 kb band
is diagnostic of homologous recombination of endogenous
20 sequences with the targeting vector (see Fig. 21a and f). Of
525 G418 and FIAU doubly-resistant clones screened by Southern
blot hybridization, 12 were found to contain the 3 kb fragment
diagnostic of recombination with the targeting vector. That
these clones represent the expected targeted events at the J_H
25 locus (as shown in Fig. 21f) was confirmed by further
digestion with HindIII, SpeI and HpaI. Hybridization of probe
A (see Fig. 21f) to Southern blots of HindIII, SpeI, and HpaI
digested DNA produces bands of 2.3 kb, >10 kb, and >10kb,
respectively, for the wild-type locus (see Fig. 21a), whereas
30 bands of 5.3 kb, 3.8 kb, and 1.9 kb, respectively, are
expected for the targeted heavy chain locus (see Fig 21f).
All 12 positive clones detected by the StuI digest showed the
predicted HindIII, SpeI, and HpaI bands diagnostic of a
targeted J_H gene. In addition, Southern blot analysis of a
35 StuI digest of all 12 clones using a neo-specific probe (probe
B, Fig. 21f) generated only the predicted fragment of 3 kb,
demonstrating that the clones each contained only a single
copy of the targeting vector.

86

Generation of mice carrying the J_H deletion

Three of the targeted ES clones described in the
previous section were thawed and injected into C57BL/6J
blastocysts as described (Bradley, A. (1987) in

5 Teratocarcinomas and Embryonic Stem Cells: A Practical
Approach, E.J. Robertson, ed. (Oxford: IRL Press), p.113-151)
and transferred into the uteri of pseudopregnant females. The
extent of ES cell contribution to the chimera was visually
estimated from the amount of agouti coat coloration, derived

10 from the ES cell line, on the black C57BL/6J background. Half
of the offspring resulting from blastocyst injection of two of
the targeted clones were chimeric (i.e., showed agouti as well
as black pigmentation); the third targeted clone did not
generate any chimeric animals. The majority of the chimeras

15 showed significant (approximately 50 percent or greater) ES
cell contribution to coat pigmentation. Since the AB-1 ES
cells are an XY cell line, most of the chimeras were male, due
to sex conversion of female embryos colonized by male ES
cells. Males chimeras were bred with C57BL/6J females and the

20 offspring monitored for the presence of the dominant agouti
coat color indicative of germline transmission of the ES
genome. Chimeras from both of the clones consistently
generated agouti offspring. Since only one copy of the heavy
chain locus was targeted in the injected ES clones, each

25 agouti pup had a 50 percent chance of inheriting the mutated
locus. Screening for the targeted gene was carried out by
Southern blot analysis of StuI-digested DNA from tail
biopsies, using the probe utilized in identifying targeted ES
clones (probe A, Fig. 21f). As expected, approximately 50

30 percent of the agouti offspring showed a hybridizing StuI band
of approximately 3 kb in addition to the wild-type band f 4.7
kb, demonstrating germline transmission of the targeted J_H
gene segment.

In order to generate mice homozygous for the

35 mutation, heterozygotes were bred together and the heavy chain
genotype of the offspring determined as described above. As
expected, three genotypes were derived from the heterozygote
matings: wild-type mice bearing two copies of the normal J_H

87

locus, heterozygotes caring one targeted copy of the gene and
one normal copy, and mice homozygous for the J_H mutation. The
absence of J_H sequences from these latter mice was verified by
hybridization of the Southern blots of StuI-digested DNA with
5   a probe specific for J_H (probe C, Fig. 21a). Whereas
hybridization of the J_H probe to a 4.7 kb fragment in DNA
samples from heterozygous and wild-type siblings was observed,
no signal was present in samples from the J_H-mutant
homozygotes, attesting to the generation of a novel mouse
10 strain in which both copies of the heavy chain gene have been
mutated by deletion of the J_H sequences.

EXAMPLE 12

Heavy Chain Minilocus Transgene

15 A.   Construction of plasmid vectors for cloning large DNA
sequences

    1.   pGP1a

        The plasmid pBR322 was digested with EcoRI and StyI
    and ligated with the following oligonucleotides:

20

    oligo-42   5'- caa gag ccc gcc taa tga gcg ggc ttt ttt ttg cat
               act gcg gcc gct -3'

    oligo-43   5'- aat tag cgg ccg cag tat gca aaa aaa agc ccg ctc
25   att agg cgg gct -3'

        The resulting plasmid, pGP1a, is designed for
    cloning very large DNA constructs that can be excised by the
30 rare cutting restriction enzyme NotI. It contains a NotI
    restriction site downstream (relative to the ampicillin
    resistance gene, AmpR) of a strong transcription termination
    signal derived from the trpA gene (Christie et al., Proc.
    Natl. Acad. Sci. USA 78:4180 (1981)). This termination signal
35 reduces the potential toxicity of coding sequences inserted
    into the NotI site by eliminating readthrough transcription
    from the AmpR gene. In addition, this plasmid is low copy
    relative to the pUC plasmids because it retains the pBR322
    copy number control region. The low copy number further
40 reduces the potential toxicity of insert sequences and reduces
    the selection against large inserts due to DNA replication.

88

The vectors pGP1b, pGP1c, pGP1d, and pGP1f are derived from pGP1a and contain different polylinker cloning sites. The polylinker sequences are given below

5

pGP1a

       NotI
     GCGGCCGC

10

pGP1b

   NotI   XhoI   ClaI   BamHI   HindIII NotI
   GCggccgcctcgagatcactatcgattaattaaggatccagcagtaagcttgcGGCCGC

15

pGI1c

   NotI   SmaI   XhoI SalI HindIII   BamHI SacII NotI
   GCggccgcatcccgggtctcgaggtcgacaagctttcgaggatccgcGGCCGC

20

pGP1d

25
   NotI   SalI HindIII ClaI BamHI XhoI   NotI
   GCggccgctgtcgacaagcttatcgatggatcctcgagtgcGGCCGC

30 pGP1f

   NotI   SalI HindIII EcoRI ClaI   KpnI BamHI XhoI   NotI
   GCggccgctgtcgacaagcttcgaattcagatcgatgtggtacctggatcctcgagtgcGGCCGC

35 Each of these plasmids can be used for the construction of large transgene inserts that are excisable with NotI so that the transgene DNA can be purified away from vector sequences prior to microinjection.

40 2.   pGP1b
            pGP1a was digested with NotI and ligated with the following oligonucleotides:

oligo-47 5'- ggc cgc aag ctt act gct gga tcc tta att aat cga
45   tag tga tct cga ggc -3'

oligo-48 5'- ggc cgc ctc gag atc act atc gat taa tta agg atc
              cag cag taa gct tgc -3'

50

89

The resulting plasmid, pGP1b, contains a short
polylinker region flanked by NotI sites. This facilitates the
construction of large inserts that can be excised by NotI
digestion.

5

3.   pGPe
The following oligonucleotides:

oligo-44   5'- ctc cag gat cca gat atc agt acc tga aac agg gct
10   tgc -3'

oligo-45   5'- ctc gag cat gca cag gac ctg gag cac aca cag cct
            tcc -3'

15
were used to amplify the immunoglobulin heavy chain 3'
enhancer (S. Petterson, et al., Nature 344:165-168 (1990))
from rat liver DNA by the polymerase chain reaction technique.
        The amplified product was digested with BamHI and
20 SphI and cloned into BamHI/SphI digested pNNO3 (pNNO3 is a pUC
derived plasmid that contains a polylinker with the following
restriction sites, listed in order: NotI, BamHI, NcoI, ClaI,
EcoRV, XbaI, SacI, XhoI, SphI, PstI, BglII, EcoRI, SmaI, KpnI,
HindIII, and NotI). The resulting plasmid, pRE3, was digested
25 with BamHI and HindIII, and the insert containing the rat Ig
heavy chain 3' enhancer cloned into BamHI/HindIII digested
pGP1b. The resulting plasmid, pGPe (Fig. 22 and Table 1),
contains several unique restriction sites into which sequences
can be cloned and subsequently excised together with the 3'
30 enhancer by NotI digestion.

TABLE 1

```
AATTAGCggccgcctcgagatcactatcgattaattaaggatccagatatcagtaccttaaacagggctttgctcataaca
ttctctcttctgtctcttctgtctctctgtgtgtgtgtgtctctctctctgtctctctgtctctctctctgtctctctgtctctgtgtgtgtt
tttctctctctgtctctctctctctgtctctctctgtctctctctgtctgtctctctgtctctctgtctctctgtctctctctctctctctctctt
ttctctctctgtctctctctcacacacacacacacacacacacacacacctgccgagtgactcactctgtgcagggcttggccc
tcggggcacatgcaaatggatgtttgttccatgcagaaaaacatgtttctcattctctgagccaaaaatagcatcaatga
tttctccaccctgcagctgcaggttcaccccacctggccaggttgaccagctttggggatggggctggggttccatgac
ccctaacggtgacattgaattcagtgtttttcccattaatcgacactgctggaatctgaccctaggagggaatgacaggag
ataggcaaggtccaaacacccagggaagtgggcagacaggaaggctgtgtgtgctccaggtccttctgcatgctgcaga
tctgaattcccgggtaccaagcttgcGGCCGCAGTATGCAAAAAAAAGCCCGCTCATTAGGCGGGCTCTTGGCAGAACAT
ATCCATCGCGTCCGCCATCTCCAGCAGCCGCACGCGGCGCATCTCGGGCAGCGTTGGGTCCTGGCCACGGGTGCGCATGA
TCGTGCTCCTGTCGTTGAGGACCCGGCTAGGCTGGCGGGGTTGCCTTACTGGTTAGCAGAATGAATCACCGATACGCGAG
CGAACGTGAAGCGACTGCTGCTGCAAAACGTCTGCTGACCTGAGCAACAACATGAATGGTCTTCGGTTTCCGTGTTTCGTA
AAGTCTGGAAACGCGGAAGTCAGCGCCCTGCACCATTATGTTCCGGATCTGCATCGCAGGATGCTGCTGGCTACCCTGTG
GAACACCTACATCTGTATTAACGAAGCGCTGGCATTGACCCTGAGTGATTTTTCTCTGGTCCCGCCGCATCCATACCGCC
AGTTGTTTACCCTCACAACGTTCCAGTAACCGGCATGTTCATCATCAGTAACCCGTATCGTGAGCATCCTCTCTCGTTT
CATCGGTATCATTACCCCCATGAACAGAAATTC  CCTTACACGGAGGCATCAAGTGACCAAACAGGAAAAACCGCCC
TAACATGGCCCGCTTTATCAGAAGCCAGACATT  GCTTCTGGAGAAACTCAACGAGCTGGACGCGGATGAACAGGCAG
ACATCTGTGAATCGCTTCACGACCACGCTGATC  TTTACCGCAGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAAC
CTCTGACACATGCAGCTCCCGGAGACGGTCACA  TGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGC
GTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCC  GACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTA
TGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACC
GCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCAC
TCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGC
CAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCT
CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCT
GTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTG
TAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCG
CCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG
ATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGT
ATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCG
CTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTT
TCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCAC
CTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT
GCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATA
ACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTT
ATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTA
ATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATCGTG
GTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTT
GTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTA
TGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCA
TTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAAC
TTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGA
TGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGG
CAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAG
CATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCA
CATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACG
AGGCCCTTTCGTCTTCAAG
```

Sequence of vector pGPe.

B.   Construction of IgM expressing minilocus transgene, pIGM1

1.   Isolation of J-μ constant region clones and construction of pJM1

A human placental genomic DNA library cloned into the phage vector λEMBL3/SP6/T7 (Clonetech Laboratories, Inc., Palo Alto, CA) was screened with the human heavy chain J region specific oligonucleotide:

oligo-1   5'- gga ctg tgt ccc tgt gtg atg ctt ttg atg tct ggg gcc aag -3'

and the phage clone λ1.3 isolated. A 6 kb HindIII/KpnI fragment from this clone, containing all six J segments as well as D segment DHQ52 and the heavy chain J-μ intronic enhancer, was isolated. The same library was screened with the human μ specific oligonucleotide:

oligo-2   5'- cac caa gtt gac ctg cct ggt cac aga cct gac cac cta tga -3'

and the phage clone λ2.1 isolated. A 10.5 kb HindIII/XhoI fragment, containing the μ switch region and all of the μ constant region exons, was isolated from this clone. These two fragments were ligated together with KpnI/XhoI digested pNNO3 to obtain the plasmid pJM1.

2.   pJM2

A 4 kb XhoI fragment was isolated from phage clone λ2.1 that contains sequences immediately downstream of the sequences in pJM1, including the so called Σμ element involved in δ-associated deleteon of the μ in certain IgD expressing B-cells (Yasui et al., Eur. J. Immunol. 19:1399 (1989), which is incorporated herein by reference). This fragment was treated with the Klenow fragment of DNA polymerase I and ligated to XhoI cut, Klenow treated, pJM1. The resulting plasmid, pJM2 (Fig. 23), had lost the internal XhoI site but retained the 3' XhoI site due to incomplete reaction by the Klenow enzyme. pJM2 contains the entire human J region, the heavy chain J-μ intronic enhancer, the μ switch region and all

of the μ constant region exons, as well as the two 0.4 kb
direct repeats, σμ and Σμ, involved in δ-associated deletion
of the μ gene.

5   3.   Isolation of D region clones and construction of pDH1
              The following human D region specific
oligonucleotide:

    oligo-4   5'- tgg tat tac tat ggt tcg ggg agt tat tat aac cac
10   agt gtc -3'

was used to screen the human placenta genomic library for D
region clones. Phage clones λ4.1 and λ4.3 were isolated. A
15 5.5 kb XhoI fragment, that includes the D elements D_{K1}, D_{N1},
and D_{M2} (Ichihara et al., EMBO J. 7:4141 (1988)), was isolated
from phage clone λ4.1. An adjacent upstream 5.2 kb XhoI
fragment, that includes the D elements D_{LR1}, D_{XP1}, $D_{XP'1}$, and
D_{A1}, was isolated from phage clone λ4.3. Each of these D
20 region XhoI fragments were cloned into the SalI site of the
plasmid vector pSP72 (Promega, Madison, WI) so as to destroy
the XhoI site linking the two sequences. The upstream
fragment was then excised with XhoI and SmaI, and the
downstream fragment with EcoRV and XhoI. The resulting
25 isolated fragments were ligated together with SalI digested
pSP72 to give the plasmid pDH1. pDH1 contains a 10.6 kb
insert that includes at least 7 D segments and can be excised
with XhoI (5') and EcoRV (3').

30 4.   pCOR1
              The plasmid pJM2 was digested with Asp718 (an
isoschizomer of KpnI) and the overhang filled in with the
Klenow fragment of DNA polymerase I. The resulting DNA was
then digested with ClaI and the insert isolated. This insert
35 was ligated to the XhoI/EcoRV insert of pDH1 and XhoI/ClaI
digested pGPe to generate pCOR1 (Fig. 24).

5.   pVH251
              A 10.3 kb genomic HindIII fragment containing the
40 two human heavy chain variable region segments V_H251 and V_H105

93

(Humphries et al., <u>Nature</u> <u>331</u>:446 (1988), which is
incorporated herein by reference) was subcloned into pSP72 to
give the plasmid pVH251.

5   6. <u>pIGM1</u>
          The plasmid pCOR1 was partially digested with XhoI
and the isolated XhoI/SalI insert of pVH251 cloned into the
upstream XhoI site to generate the plasmid pIGM1 (Fig. 25).
pIGM1 contains 2 functional human variable region segments, at
10   least 8 human D segments all 6 human J_H segments, the human
$J-\mu$ enhancer, the human $\sigma\mu$ element, the human μ switch region,
all of the human μ coding exons, and the human $\Sigma\mu$ element,
together with the rat heavy chain 3' enhancer, such that all
of these sequence elements can be isolated on a single
15   fragment, away from vector sequences, by digestion with NotI
and microinjected into mouse embryo pronuclei to generate
transgenic animals.

C.   <u>Construction of IgM and IgG expressing minilocus</u>
20   <u>transgene, pHC1</u>
     1.   <u>Isolation of γ constant region clones</u>
          The following oligonucleotide, specific for human Ig
g constant region genes:

25   oligo-29   5'- cag cag gtg cac acc caa tgc cca tga gcc cag aca
                ctg gac -3'

was used to screen the human genomic library. Phage clones
129.4 and λ29.5 were isolated. A 4 kb HindIII fragment of
30   phage clone λ29.4, containing a γ switch region, was used to
probe a human placenta genomic DNA library cloned into the
phage vector lambda FIX™ II (Stratagene, La Jolla, CA). Phage
clone λSg1.13 was isolated. To determine the subclass of the
different γ clones, dideoxy sequencing reactions were carried
35   out using subclones of each of the three phage clones as
templates and the following oligonucleotide as a primer:

oligo-67   5'- tga gcc cag aca ctg gac -3'

94

Phage clones λ29.5 and λSγ1.13 were both determined to be of the γ1 subclass.

2. pγe1

5   A 7.8 kb HindIII fragment of phage clone λ29.5, containing the γ1 coding region was cloned into pUC18. The resulting plasmid, pLT1, was digested with XhoI, Klenow treated, and religated to destroy the internal XhoI site. The resulting clone, pLT1xk, was digested with HindIII and the

10 insert isolated and cloned into pSP72 to generate the plasmid clone pLT1xks. Digestion of pLT1xks at a polylinker XhoI site and a human sequence derived BamHI site generates a 7.6 kb fragment containing the γ1 constant region coding exons. This 7.6 kb XhoI/BamHI fragment was cloned together with an

15 adjacent downstream 4.5 kb BamHI fragment from phage clone λ29.5 into XhoI/BamHI digested pGPe to generate the plasmid clone pγe1. pγe1 contains all of the γ1 constant region coding exons, together with 5 kb of downstream sequences, linked to the rat heavy chain 3' enhancer.

20

3. pγe2

        A 5.3 kb HindIII fragment containing the γ1 switch region and the first exon of the pre-switch sterile transcript (P. Sideras et al. (1989) International Immunol. 1, 631) was

25 isolated from phage clone λSγ1.13 and cloned into pSP72 with the polylinker XhoI site adjacent to the 5' end of the insert, to generate the plasmid clone pSγ1s. The XhoI/SalI insert of pSγ1s was cloned into XhoI digested pγe1 to generate the plasmid clone pγe2 (Fig. 26). pγe2 contains all of the γ1

30 constant region coding exons, and the upstream switch region and sterile transcript exons, together with 5 kb of downstream sequences, linked to the rat heavy chain 3' enhancer. This clone contains a unique XhoI site at the 5' end of the insert. The entire insert, together with the XhoI site and the 3' rat

35 enhancer can be excised from vector sequences by digestion with NotI.

95

4.   pHC1

         The plasmid pIGM1 was digested with XhoI and the 43
kb insert isolated and cloned into XhoI digested pge2 to
generate the plasmid pHC1 (Fig. 25).   pHC1 contains 2
5 functional human variable region segments, at least 8 human D
segments all 6 human J_H segments, the human $J-\mu$ enhancer, the
human $\sigma\mu$ element, the human μ switch region, all of the human
μ coding exons, the human $\Sigma\mu$ element, and the human γ1
constant region, including the associated switch region and
10 sterile transcript associated exons, together with the rat
heavy chain 3' enhancer, such that all of these sequence
elements can be isolated on a single fragment, away from
vector sequences, by digestion with NotI and microinjected
into mouse embryo pronuclei to generate transgenic animals.

15

D.   Construction of IgM and IgG expressing minilocus
transgene, pHC2
1.   Isolation of human heavy chain V region gene VH49.8
         The human placental genomic DNA library lambda, FIX™
20 II, Stratagene, La Jolla, CA) was screened with the following
human VH1 family specific oligonucleotide:

oligo-49   5'- gtt aaa gag gat ttt att cac ccc tgt gtc ctc tcc
            aca ggt gtc -3'
25

         Phage clone λ49.8 was isolated and a 6.1 kb XbaI
fragment containing the variable segment VH49.8 subcloned into
pNNO3 (such that the polylinker ClaI site is downstream of
30 VH49.8 and the polylinker XhoI site is upstream) to generate
the plasmid pVH49.8. An 800 bp region of this insert was
sequenced, and VH49.8 found to have an open reading frame and
intact splicing and recombination signals, thus indicating
that the gene is functional (Table 2).

TABLE 2

```
TTCCTCAGGC AGGATTTAGG GCTTGGTCTC TCAGCATCCC ACACTTGTAC      50

AGCTCTTGTG GCATCTCTGT TTTCTTTCTC ATCCTAGTTC AAGCTTTGTG     100

CTGTGTAATA CCCTGCCTCA TGAATATGCA AATAATCTGA GGTCTTCTGA     150

CATAAATATA GATATTTTGG TCCCCTGTGA GCATCACATA ACAACCAGAT     200

TCCTCCTCTA AAGAGCCCC  TGGGAGCACA GCTCATCACC ATGGACTGGA     250
                                                MetAspTrpT

CCTGGAGGTT CCTCTTTGTG GTGGCAGCAG CTACAGgtaa ggggcttcct     300
hrTrpArgPh eLeuPheVal ValAlaAlaA laThr

agtcctaagg ctgaggaagg gatcctggtt tagttaaaga ggattttatt     350

cacccct     tcccctccac agGTGTCCAG TCCCAGGTCC AGCTGGTGCA     400
                         GlyValGln  SerGlnValG lnLeuValGl

GTCTGGGGCT GAGGTGAAGA AGCCTGGGTC CTCGGTGAAG GTCTCCTGCA     450
nSerGlyAla GluValLysL ysProGlySe rSerValLys ValSerCysL

AGGCTTCTGG AGGCACCTTC AGCAGCTATG CTATCAGCTG GGTGCGACAG     500
ysAlaSerGl yGlyThrPhe SerSerTyrA laIleSerTr pValArgGln

GCCCCTGGAC AAGGGCTTGA GTGGATGGGA AGGATCATCC CTATCCTTGG     550
AlaProGlyG lnGlyLeuGl uTrpMetGly ArgIleIleP roIleLeuGl

TATAGCAAAC TACGCACAGA AGTTCCAGGG CAGAGTCACG ATTACCGCGG     600
yIleAlaAsn TyrAlaGlnL ysPheGlnGl yArgValThr IleThrAlaA

ACAAATCCAC GAGCACAGCC TACATGGAGC TGAGCAGCCT GAGATCTGAG     650
spLysSerTh rSerThrAla TyrMetGluL euSerSerLe uArgSerGlu

GACACGGCCG TGTATTACTG TGCGAGAGAT ACAGTCTTGA AACCCACATC     700
AspThrAlaV alTyrTyrCy sAlaArg

CTGAGAGTCT CAGAAACCCT GAGGGAGAAG GCAGCTGTGC CGGGCTGAGG     750

AGATGACAGG GTTTATTAGG TTTAAGGCTG TTTACAAAAT GGGTTATATA     800

                                                         812
TTTGAGAAAA AA
```

Sequence of human $V_H I$ family gene $V_H 49.8$

97

2.   pV2

A 4 kb XbaI genomic fragment containing the human
V_HIV family gene V_H4-21 (Sanz et al., <u>EMBO J</u>., <u>8</u>:3741 (1989)),
subcloned into the plasmid pUC12, was excised with SmaI and
HindIII, and treated with the Klenow fragment of polymerase I.
The blunt ended fragment was then cloned into ClaI digested,
Klenow treated, pVH49.8. The resulting plasmid, pV2, contains
the human heavy chain gene VH49.8 linked upstream of VH4-21 in
the same orientation, with a unique SalI site at the 3' end of
the insert and a unique XhoI site at the 5' end.

3.   pSγ1-5'

A 0.7 kb XbaI/HindIII fragment (representing
sequences immediately upstream of, and adjacent to, the 5.3 kb
γ1 switch region containing fragment in the plasmid pγe2)
together with the neighboring upstream 3.1 kb XbaI fragment
were isolated from the phage clone λSg1.13 and cloned into
HindIII/XbaI digested pUC18 vector. The resulting plasmid,
pSγ1-5', contains a 3.8 kb insert representing sequences
upstream of the initiation site of the sterile transcript
found in B-cells prior to switching to the γ1 isotype (P.
Sideras et al., <u>International Immunol.</u> 1:631 (1989)). Because
the transcript is implicated in the initiation of isotype
switching, and upstream cis-acting sequences are often
important for transcription regulation, these sequences are
included in transgene constructs to promote correct expression
of the sterile transcript and the associated switch
recombination.

4.   pVGE1

The pSγ1-5' insert was excised with SmaI and
HindIII, treated with Klenow enzyme, and ligated with the
following oligonucleotide linker:

5'- ccg gtc gac cgg -3'

The ligation product was digested with SalI and ligated to
SalI digested pV2. The resulting plasmid, pVP, contains 3.8

kb of γ1 switch 5' flanking sequences linked downstream of the
two human variable gene segments VH49.8 and VH4-21 (see Table
2). The pVP insert is isolated by partial digestion with SalI
and complete digestion with XhoI, followed by purification of

5   the 15 kb fragment on an agarose gel. The insert is then
cloned into the XhoI site of pγe2 to generate the plasmid
clone pVGE1 (Fig. 27). pVGE1 contains two human heavy chain
variable gene segments upstream of the human γ1 constant gene
and associated switch region. A unique SalI site between the

10 variable and constant regions can be used to clone in D, J,
and μ gene segments. The rat heavy chain 3' enhancer is
linked to the 3' end of the γ1 gene and the entire insert is
flanked by NotI sites.

15 5.   pHC2

       The plasmid clone pVGE1 is digested with SalI and
the XhoI insert of pIGM1 is cloned into it. The resulting
clone, pHC2 (Fig. 25), contains 4 functional human variable
region segments, at least 8 human D segments all 6 human J_H

20 segments, the human J-m enhancer, the human σμ element, the
human μ switch region, all of the human μ coding exons, the
human Σμ element, and the human γ1 constant region, including
the associated switch region and sterile transcript associated
exons, together with 4 kb flanking sequences upstream of the

25 sterile transcript initiation site. These human sequences are
linked to the rat heavy chain 3' enhancer, such that all of
the sequence elements can be isolated on a single fragment,
away from vector sequences, by digestion with NotI and
microinjected into mouse embryo pronuclei to generate

30 transgenic animals. A unique XhoI site at the 5' end of the
insert can be used to clone in additional human variable gene
segments to further expand the recombinational diversity of
this heavy chain minilocus.

35 E.   Transgenic mice

       The NotI inserts of plasmids pIGM1 and pHC1 were
isolated from vector sequences by agarose gel electrophoresis.
The purified inserts were microinjected into the pronuclei of

99

fertilized (C57BL/6 x CBA)F2 mouse embryos and transferred the
surviving embryos into pseudopregnant females as described by
Hogan et al. (B. Hogan, F. Costantini, and E. Lacy, Methods of
Manipulating the Mouse Embryo, 1986, Cold Spring Harbor

5   Laboratory, New York). Mice that developed from injected
embryos were analyzed for the presence of transgene sequences
by Southern blot analysis of tail DNA. Transgene copy number
was estimated by band intensity relative to control standards
containing known quantities of cloned DNA. At 3 to 8 weeks of

10   age, serum was isolated from these animals and assayed for the
presence of transgene encoded human IgM and IgG1 by ELISA as
described by Harlow and Lane (E. Harlow and D. Lane.
Antibodies: A Laboratory Manual, 1988, Cold Spring Harbor
Laboratory, New York). Microtiter plate wells were coated

15   with mouse monoclonal antibodies specific for human IgM (clone
AF6, #0285, AMAC, Inc. Westbrook, ME) and human IgG1 (clone
JL512, #0280, AMAC, Inc. Westbrook, ME). Serum samples were
serially diluted into the wells and the presence of specific
immunoglobulins detected with affinity isolated alkaline

20   phosphatase conjugated goat anti-human Ig (polyvalent) that
had been pre-adsorbed to minimize cross-reactivity with mouse
immunoglobulins. Table 3 and Fig. 28 show the results of an
ELISA assay for the presence of human IgM and IgG1 in the
serum of two animals that developed from embryos injected with

25   the transgene insert of plasmid pHC1. All of the control non-
transgenic mice tested negative for expression of human IgM
and IgG1 by this assay. Mice from two lines containing the
pIGM1 NotI insert (lines #6 and 15) express human IgM but not
human IgG1. We tested mice from 6 lines that contain the pHC1

30   insert and found that 4 of the lines (lines #26, 38, 57 and
122) express both human IgM and human IgG1, while mice from
two of the lines (lines #19 and 21) do not express detectable
levels of human immunoglobulins. The pHC1 transgenic mice
that did not express human immunoglobulins were so-called G_o

35   mice that developed directly from microinjected embryos and
may have been mosaic for the presence of the transgene.
Southern blot analysis indicates that many of these mice
contain one or fewer copies of the transgene per cell. The

detection of human IgM in the serum of pIGM1 transgenics, and
human IgM and IgG1 in pHC1 transgenics, provides evidence that
the transgene sequences function correctly in directing VDJ
joining, transcription, and isotype switching. One of the
5 animals (#18) was negative for the transgene by Southern blot
analysis, and showed no detectable levels of human IgM or
IgG1. The second animal (#38) contained approximately 5
copies of the transgene, as assayed by Southern blotting, and
showed detectable levels of both human IgM and IgG1. The
10 results of ELISA assays for 11 animals that developed from
transgene injected embryos is summarized in the table below
(Table 3).

TABLE 3

15 Detection of human IgM and IgG1 in the serum of transgenic
animals by ELISA assay

animal #	injected transgene	approximate transgene copies per cell	human IgM	human IgG1
6	pIGM1	1	++	-
7	pIGM1	0	-	-
9	pIGM1	0	-	-
10	pIGM1	0	-	-
12	pIGM1	0	-	-
15	pIGM1	10	++	-
18	pHC1	0	-	-
19	pHC1	1	-	-
21	pHC1	<1	-	-
26	pHC1	2	++	+
38	pHC1	5	++	+

101

Table 3 shows a correlation between the presence of
integrated transgene DNA and the presence of transgene encoded
immunoglobulins in the serum. Two of the animals that were
5   found to contain the pHC1 transgene did not express detectable
levels of human immunoglobulins. These were both low copy
animals and may not have contained complete copies of the
transgenes, or the animals may have been genetic mosaics
(indicated by the <1 copy per cell estimated for animal #21),
10   and the transgene containing cells may not have populated the
hematopoietic lineage. Alternatively, the transgenes may have
integrated into genomic locations that are not conducive to
their expression. The detection of human IgM in the serum of
pIGM1 transgenics, and human IgM and IgG1 in pHC1 transgenics,
15   indicates that the transgene sequences function correctly in
directing VDJ joining, transcription, and isotype switching.

F.   cDNA clones

        To assess the functionality of the pHC1 transgene in
20   VDJ joining and class switching, as well the participation of
the transgene encoded human B-cell receptor in B-cell
development and allelic exclusion, the structure of
immunoglobulin cDNA clones derived from transgenic mouse
spleen mRNA were examined. The overall diversity of the
25   transgene encoded heavy chains, focusing on D and J segment
usage, N region addition, CDR3 length distribution, and the
frequency of joints resulting in functional mRNA molecules was
examined. Transcripts encoding IgM and IgG incorporating
VH105 and VH251 were examined.

30   Polyadenylated RNA was isolated from an eleven week
old male second generation line-57 pHC1 transgenic mouse.
This RNA was used to synthesize oligo-dT primed single
stranded cDNA. The resulting cDNA was then used as template
for four individual PCR amplifications using the following
35   four synthetic oligonucleotides as primers: VH251 specific
oligo-149, cta gct cga gtc caa gga gtc tgt gcc gag gtg cag ctg
(g,a,t,c); VH105 specific o-150, gtt gct cga gtg aaa ggt gtc
cag tgt gag gtg cag ctg (g,a,t,c); human gamma1 specific
oligo-151, ggc gct cga gtt cca cga cac cgt cac cgg ttc; and

102

human mu specific oligo-152, cct gct cga ggc agc caa cgg cca
cgc tgc tcg. Reaction 1 used primers o-149 and o-151 to
amplify VH251-gamma1 transcripts, reaction 2 used o-149 and o-
152 to amplify VH251-mu transcripts, reaction 3 used o-150 and

5   o-151 to amplify VH105-gamma1 transcripts, and reaction 4 used
o-150 and o-152 to amplify VH105-mu transcripts. The
resulting 0.5 kb PCR products were isolated from an agarose
gel; the μ transcript products were more abundant than the γ
transcript products, consistent with the corresponding ELISA

10 data (Fig. 34). The PCR products were digested with XhoI and
cloned into the plasmid pNN03. Double-stranded plasmid DNA
was isolated from minipreps of nine clones from each of the
four PCR amplifications and dideoxy sequencing reactions were
performed. Two of the clones turned out to be deletions

15 containing no D or J segments. These could not have been
derived from normal RNA splicing products and are likely to
have originated from deletions introduced during PCR
amplification. One of the DNA samples turned out to be a
mixture of two individual clones, and three additional clones

20 did not produce readable DNA sequence (presumably because the
DNA samples were not clean enough). The DNA sequences of the
VDJ joints from the remaining 30 clones are compiled in Table
4. Each of the sequences are unique, indicating that no
single pathway of gene rearrangement, or single clone of

25 transgene expressing B-cells is dominant. The fact that no
two sequences are alike is also an indication of the large
diversity of immunoglobulins that can be expressed from a
compact minilocus containing only 2 V segments, 10 D segments,
and 6 J segments. Both of the V segments, all six of the J

30 segments, and 7 of the 10 D segments that are included in the
transgene are used in VDJ joints. In addition, both constant
region genes (mu and gamma1) are incorporated into
transcripts. The VH105 primer turned out not to be specific
for VH105 in the reactions performed. Therefore many of the

35 clones from reactions 3 and 4 contained VH251 transcripts.
Additionally, clones isolated from ligated reaction 3 PCR
product turned out to encode IgM rather than IgG; however this
may reflect contamination with PCR product from reaction 4 as

103

the DNA was isolated on the same gel. An analogous
experiment, in which immunoglobulin heavy chain sequences were
amplified from adult human peripheral blood lymphocytes (PBL),
and the DNA sequence of the VDJ joints determined, was
recently reported by Yamada et al. (J. Exp. Med. 173:395-407
(1991), which is incorporated herein by reference). We
compared the data from human PBL with our data from the pHC1
transgenic mouse.

TABLE 4

		V	n-D-n	J	C
1	VH251 DHQ52 J3 ψ	TACTGTGCGAGA	GCGGTAACTGACATCGAY	GCTTTTGATATCTGGGGCC AAGGGACCAATGGTCACCGTCTCCTCAG	CTTCA ACCAAG
2	VH251 DN1 J4 ψ	TACTGTGCGAGA	CACGGTATAGCAGCAGCTGG	CTTTGACTACTGGGGCC AGGGAACCCTGGTCACCGTCTCCTCAG	CCTCCACCAAG
3	VH251 D7 J6 ψ	TACTGTGCGAGA	T	ATTACTACTACTACTACGGTATGGACGTCTGGGGCC AAGGGACCACGGTCACCGTCTCCTCAG	CCTCCACCAAG
4	VH251 DXP'1 J6 ψ	TACTGTGCGAGA	CATTACGATATTTGACTGGTG	CTACTACTACTACTACGGTATGGACGTCTGGGGCC AGGGACCACGGTCACCGTCTCCTCAG	CCTCCACCAAG
5	VH251 DXP'1 J4 ψ	TACTGTGCGAGA	DCGGGTACTATGGTTCGGGCGAGTATAAACGT	CTTTGACTACTGGGGCC CAAGGGACAATGGTCACCGTCTCCTCAG	CCTCCACCAAG
6	VH251 D7 J3 ψ	TACTGTGCGAGA	GGGTGTGGCGAT	GTTTTGATATCTGGGGCC CAAKCAACAATGGTCACCGTCTCCTCAG	GGAGTGCATCC
7	VH251 DHQ52 J3 ψ	TACTGTGCGAGA	GCAACTGCC	GGTTTGATATCTGGGGCC CAAKCAACAATGGTCACCGTCTCCTCAG	GGAGTGCATCC
8	VH251 DHQ52 J6 ψ	TACTGTGCGAGA	TDRGCTAACTGCGGAYC	CTACTACTACTACGGTATGGACGTCTGGGGCC CAAKCACCGTCACCGTCTCCTCAG	GGAGTGCATCC
9	VH251 - J1 ψ	TACTGTGCGAGA		TACTCCAGCACTGGGGA TACKACACCCTGGTCACCGTCTCCTCAG	GGAGTGCATCC
10	VH251 D1R2 J4 ψ	TACTGTGCGAGA	CACGTACTGACTCC	TTTGACTACTGGGGCCAAGGGAACCCTGGTCACCGTCTCCTCAG	GGAGTGCATCC
11	VH251 DXP'1 J4 ψ	TACTGTGCGAGA	CAGATTACTATGGTTCGGGGAGTTGC	CTTTGACTACTGGGGCCCTAAGGGAACCCTGGTCACCGTCTCCTCAG	GGAGTGCATCC
12	VH251 D7 J1 ψ	TACTGTGCGAGA	C	AATACTTCCAGCACTGGGGCCAGGGCACCCTGGTCACCGTCTCCTCAG	GGAGTGCATCC
13	VH251 DHQ52 J5 ψ	TACTGTGCGAGA	CAAACTGCGG	ACTACTACTACTACCGGTATGGACGTCTGGGGCCAAGGACCACGGTCACCGTCTCCTCAG	GGAGTGCATCC
14	VH251 DXP'1 J6 ψ	TACTGTGCGAGA	CATTACTATGGTTCGGGGAGTTATG	ACTACTACTACTACTACGGTATGGAGGTCTGGGGCCCAAGGGACCACGGTCACCGTCTCCTCAG	CCTCCACCAAG
15	VH251 DXP'1 J4 ψ	TACTGTGCGAGA	CGCGGAG	TGGGGCCCAGGGAACCCTGGTCACCGTCACCGTCTCCTCAG	GGAGTGCATCC
16	VH105 DXP'1 J5 ψ	TACTGTGTGAGA	TTCTGCGGAG	ACTGGTTCGACCCCTGGGCCTGGGCCACCGGAACCCTGGTCACCGTCTCCTCAG	CCTCCACCAAG
17	VH251 DXP'1 J4 ψ	TACTGTGCGAGA	DCGGGTACTATGGTTCGGGGAGTTATTATAACGT	CTTTGACTACTACTGGGGCCCAGGGACCACCTGGTCACCGTCTCCTCAG	CCTCCACCAAG
18	VH251 DHQ52 J4 ψ	TACTGTGCGAGA	CAAACTGCGGAGA	GACTACTGGGGCCACGGAACCCTGGTCACCGTCTCCTCAG	CCTCCACCAAG
19	VH251 DN1 J6 ψ	TACTGTGCGAGA	CGATATAGTGCCTACGGAT	ACTACTACGGTATGGACGTCTGGGGCCCAAGGGACCACCGTCACCGTCTCCTCAG	CCTCCACCAAG
20	VH251 DHQ52 J6 ψ	TACTGTGCGAGA	CAAACTGCGGAG	ACTACTTTGACTACTACTGGGGCCCAAGAACCCTGGTCACCGTCTCCTCAG	CGAGTGCATCC
21	VH251 DN1 J2 ψ	TACTGTGCGAGA	TATAGTGCCTACGGTAC	CTACTGGGGACTTTCGGATCTCTGGGGCCGTGCACCGTCACTGTGTCCTCAG	CCTCCACCAAG
22	VH251 DIR2 J6 ψ	TACTGTGCGAGA	GCATCCCCCCCTCCTTT	ACTACTACGGTATGGACGTCTGGCGTCGGCTCAAGGGACCACGGTCACCGTCTCCTCAG	CCTCCACCAAG
23	VH251 DN1 J4 ψ	TACTGTGCGAGA	CGCCGGTGGGGC	TTTGACTACTGGGGCCCAGGAACCCTGGTCACCGTCTCCTCAG	GGAGTGCATCC
24	VH105 D7 J6 ψ	TACTGTGTG	CCCGCGGAACT	TTACTACTACTACTACGGTATGGACGTCTGGGGCCCAAGGACCACCGTCACCGTCTCCTCAG	GGAGTGCATCC
25	VH105 DXP1 J4 ψ	TACTGTGTGAGA	GATATTTTGACTGGTTAGG	TGACTACTGGGGCC CACCAACCCTGGTCACCGTCTCCTCAG	GGAGTGCATCC
26	VH251 DN1 J3 ψ	TACGTGCGAGA	CATCGTATACCAGCAGCTGGTAC	TGCTTTTGATATACCAGCACCAGGGACAATGGTCACCGTCTTCAG	GGAGTGCATCC
27	VH105 DHQ52 J3 ψ	TACTGTGTGAGA	TCTACTGCGGTTG	ATCGTTTGATATCTGGGCGCCAAGGGACCAATGGTCACCGTCTCTTCAG	GGAGTGCATCC
28	VH251 DN1 J4 ψ	TACTGTGCG	CABATAGCCACACCTGCC	CTACTTTGACTACTGGGGCCCAAGGGACCGTCGGTGTCACCGTCTCCTCAG	GGAGTGCATCC
29	VH105 DN1 J4 ψ	TACTGTGTG	TGTATAGCCACCTCGTATAGGGAGCGA	CTACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCAG	GGAGTGCATCC
30	VH251 DHQ52 J4 ψ	TACTGTGCGAGA	CAAACTGCGG	TGACTACTGGGGCGGACTGGCGACGGAACCCTGCTGGTGTCACCGTCTCCTCAG	GGAGTGCATCC

105

G. J segment choice

          Table 5 compared the distribution of J segments
incorporated into pHC1 transgene encoded transcripts to J
segments found in adult human PBL immunoglobulin transcripts.
5 The distribution profiles are very similar, J4 is the dominant
segment in both systems, followed by J6. J2 is the least
common segment in human PBL and the transgenic animal.

TABLE 5   J. Segment Choice
10
                              Percent Usage (± 3%)
          J. Segment   HC1 transgenic   Human PBL

            J1   7   1
15   J2   3   <1
            J3   17   9
            J4   44   53
            J5   3   15
            J6   26   22
20
                                  100%   100%

H. D segment choice

25   49% (40 of 82) of the clones analyzed by Yamada et
al. incorporated D segments that are included in the pHC1
transgene. An additional 11 clones contained sequences that
were not assigned by the authors to any of the known D
segments. Two of these 11 unassigned clones appear to be
30 derived from an inversion of the DIR2 segments which is
included in the pHC1 construct. This mechanism, which was
predicted by Ichihara et al. (EMBO J. 7:4141 (1988)) and
observed by Sanz (J. Immunol. 147:1720-1729 (1991)), was not
considered by Yamada et al. (J. Exp. Med. 173:395-407 (1991)).
35 Table 5 is a comparison of the D segment distribution for the
pHC1 transgenic mouse and that observed for human PBL
transcripts by Yamada et al. The data of Yamada et al. was
recompiled to include DIR2 use, and to exclude D segments that
are not in the pHC1 transgene. Table 6 demonstrates that the
40 distribution of D segment incorporation is very similar in the
transgenic mouse and in human PBL. The two dominant human D
segments, DXP'1 and DN1, are also found with high frequency in
the transgenic mouse. The most dramatic dissimilarity between

106

the two distributions is the high frequency of DHQ52 in the transgenic mouse as compared to the human. The high frequency of DHQ52 is reminiscent of the D segment distribution in the human fetal liver. Sanz has observed that 14% of the heavy

5 chain transcripts contained DHQ52 sequences. If D segments not found in pHC1 are excluded from the analysis, 31% of the fetal transcripts analyzed by Sanz contain DHQ52. This is comparable to the 27% that we observe in the pHC1 transgenic mouse.

10

TABLE 6   D Segment Choice

D. Segment	Percent Usage (± 3%)	
	HC1 transgenic	Human PBL
DLR1	<1	<1
DXP1	3	6
DXP'1	25	19
DA1	<1	12
DK1	7	12
DN1	12	22
DIR2	7	4
DM2	<1	2
DLR2	3	4
DHQ52	26	2
?	17	17
	100%	100%

30 I. Functionality of VDJ joints

Table 7 shows the predicted amino acid sequences of the VDJ regions from 30 clones that were analyzed from the pHC1 transgenic. The translated sequences indicate that 23 of the 30 VDJ joints (77%) are in-frame with respect to the

35 variable and J segments.

TABLE 7   Functionality of V-D-J Joints

				FR3	CDR3	FR4
1	VH251	DHQ52	J3	γ1	YCAR	FLTGVDAFDI	WGQGTMVTVSSASTK
2	VH251	DN1	J4	γ1	YCAR	HRIAAAGFDY	WGQGTLVTVSSASTK
3	VH251	D?	J6	γ1	YCAR	YYYYYGMDV	WGQGTTVTVSSASTK
4	VH251	DXP'1	J6	γ1	YCAR	HYDILTGPTTTVWTSGAKGPRSPSPQPPP	
5	VH251	DXP'1	J4	γ1	YCAR	RRYYGSGSYYNVFDY	WGQGTLVTVSSASTK
6	VH251	D?	J3	γ1	YCAR	RGVSDAFDI	WGQGTMVTVSSASTK
7	VH251	DHQ52	J3	μ	YCAR	ATGAFDI	WGQGDMVTVSSGSAS
8	VH251	DHQ52	J6	μ	YCAR	SAMMGSYYYYGMDV	WGQGTTVTVSSGSAS
9	VH251	-	J1	μ	YCAR	YFQH	WGQGTLVTVSSGSAS
10	VH251	DLR2	J4	μ	YCAR	HVANSFDY	WGQGTLVTVSSGSAS
11	VH251	DXP'1	J4	μ	YCAR	QITMVRGVPFDY	WGQGTLVTVSSGSAS
12	VH251	D?	J1	μ	YCAR	QYFQH	WGQGTLVTVSSGSAS
13	VH251	DHQ52	J6	μ	YCAR	QTGDYYYGMDV	WGQGTTVTVSSGSAS
14	VH251	DXP'1	J6	μ	YCAR	HYYGSGSYDYYYGMDV	WGQGTTVTVSSGSAS
15	VH251	DXP'1	J4	γ1	YCAR	QGVGPGHPGHRLLSLEQ	
16	VH105	DXP'1	J3	μ	YCVR	FWIIGHPGAKEPWSPSPQGVH	
17	VH251	DXP'1	J4	γ1	YCAR	RRYYGSGSYYNVFDY	WGQGTLVTVSSASTK
18	VH251	DHQ52	J4	γ1	YCAR	QTMGGDY	WGQGTLVTVSSASTK
19	VH251	DK1	J6	γ1	YCAR	GISGYDHYYYGIHV	WGQGTTVTVSSASTK
20	VH251	DHQ52	J4	μ	YCAR	QTGEDYFDY	WGQGTLVTVSSGSAS
21	VH251	DK1	J2	γ1	YCAR	YSGYDYLLVLRSLGPWHPGHCLLSLHR	
22	VH251	DIR2	J6	γ1	YCAR	AELPSFDYYGMDV	WGQGTTVTVSSASTK
23	VH251	DIR2	J4	μ	YCAR	RGGLYTGAKEPWSPSPQGVH	
24	VH105	D?	J6	μ	YCVP	VETILLLRYGHLGPRDHGHRLLRDCI	
25	VH105	DXP1	J4	μ	YCVR	DILTGRDY	WGQGTLVTVSSGSAS
26	VH251	DK1	J3	μ	YCAR	HRIAAAGTAFDI	WGQGDMVTVSSGSAS
27	VH105	DHQ52	J3	μ	YCVR	STGVDAFDI	WGQGDMVTVSSGSAS
28	VH251	DK1	J4	μ	YCAR	IAAAILELLGPGHPGHRLLRDCI	
29	VH105	DK1	J4	μ	YCVC	IAAAGHHHY	WGQGDLVTVSSGSAS
30	VH251	DHQ52	J4	μ	YCAR	QMGDY	WGQGDLVTVSSGSAS

108

J.   CDR3 length distribution

Table 8 compared the length of the CDR3 peptides
from transcripts with in-frame VDJ joints in the pHC1
transgenic mouse to those in human PBL. Again the human PBL
5   data comes from Yamada et al. The profiles are similar with
the transgenic profile skewed slightly toward smaller CDR3
peptides than observed from human PBL. The average length of
CDR3 in the transgenic mouse is 10.3 amino acids. This is
substantially the same as the average size reported for
10 authentic human CDR3 peptides by Sanz (J. Immunol. 147:1720-
1729 (1991)).

TABLE 8   CDR3 Length Distribution

#amino acids in CDR3	Percent Occurrence (± 3%)	
	HC1 transgenic	Human PBL
3-8	26	14
9-12	48	41
13-18	26	37
19-23	<1	7
>23	<1	1
	100%	100%

EXAMPLE 13

30 Rearranged Heavy Chain Transgenes
A.   Isolation of Rearranged Human Heavy Chain VDJ segments.
Two human leukocyte genomic DNA libraries cloned
into the phage vector λEMBL3/SP6/T7 (Clonetech Laboratories,
Inc., Palo Alto, CA) are screened with a 1 kb PacI/HindIII
35 fragment of λ1.3 containing the human heavy chain J-μ intronic
enhancer. Positive clones are tested for hybridization with a
mixture of the following V_H specific oligonucleotides:

oligo-7   5'-tca gtg aag gtt tcc tgc aag gca tct gga tac acc
40   ttc acc-3'

oligo-8   5'-tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc
ttc ac '

Clones that hybridized with both V and J-μ probes
are isolated and the DNA sequence of the rearranged VDJ
segment determined.

5 B. Construction of rearranged human heavy chain transgenes

Fragments containing functional VJ segments (open
reading frame and splice signals) are subcloned into the
plasmid vector pSP72 such that the plasmid derived XhoI site
is adjacent to the 5' end of the insert sequence. A subclone
10 containing a functional VDJ segment is digested with XhoI and
PacI (PacI, a rare-cutting enzyme, recognizes a site near the
J-m intronic enhancer), and the insert cloned into XhoI/PacI
digested pHC2 to generate a transgene construct with a
functional VDJ segment, the J-μ intronic enhancer, the μ
15 switch element, the μ constant region coding exons, and the γ1
constant region, including the sterile transcript associated
sequences, the γ1 switch, and the coding exons. This
transgene construct is excised with NotI and microinjected
into the pronuclei of mouse embryos to generate transgenic
20 animals as described above.

EXAMPLE 14

Light Chain Transgenes
A. Construction of Plasmid vectors
25 1. Plasmid vector pGP1c

Plasmid vector pGP1a is digested with NotI and the
following oligonucleotides ligated in:

oligo-81   5'-ggc cgc atc ccg ggt ctc gag gtc gac aag ctt tcg
30   agg atc cgc-3'

oligo-82   5'-ggc cgc gga tcc tcg aaa gct tgt cga cct cga gac
           ccg gga tgc-3'

35 The resulting plasmid, pGP1c, contains a polylinker with XmaI,
XhoI, SalI, HindIII, and BamHI restriction sites flanked by
NotI sites.

110

2. Plasmid vector pGP1d

Plasmid vector pGP1a is digested with NotI and the following oligonucleotides ligated in:

5 oligo-87 5'-ggc cgc tgt cga caa gct tat cga tgg atc ctc gag
              tgc -3'

   oligo-88 5'-ggc cgc act cga gga tcc atc gat aag ctt gtc gac
              agc -3'

10

The resulting plasmid, pGP1d, contains a polylinker with SalI, HindIII, ClaI, BamHI, and XhoI restriction sites flanked by NotI sites.

15 B. Isolation of Jκ and Cκ clones

A human placental genomic DNA library cloned into the phage vector λEMBL3/SP6/T7 (Clonetech Laboratories, Inc., Palo Alto, CA) was screened with the human kappa light chain J region specific oligonucleotide:

20

   oligo-36 5'- cac ctt cgg cca agg gac acg act gga gat taa acg
              taa gca -3'

and the phage clones 136.2 and 136.5 isolated. A 7.4 kb XhoI
25 fragment that includes the Jκ1 segment was isolated from
136.2 and subcloned into the plasmid pNNO3 to generate the
plasmid clone p36.2. A neighboring 13 kb XhoI fragment that
includes Jk segments 2 through 5 together with the Cκ gene
segment was isolated from phage clone 136.5 and subcloned into
30 the plasmid pNNO3 to generate the plasmid clone p36.5.
Together these two clones span the region beginning 7.2 kb
upstream of Jκ1 and ending 9 kb downstream of Cκ.

C. Construction of rearranged light chain transgenes

35 1. pCK1, a Cκ vector for expressing rearranged variable
segments

The 13 kb XhoI insert of plasmid clone p36.5
containing the Cκ gene, together with 9 kb of downstream

111

sequences, is cloned into the SalI site of plasmid vector
pGP1c with the 5' end of the insert adjacent to the plasmid
XhoI site. The resulting clone, pCK1 can accept cloned
fragments containing rearranged VJκ segments into the unique
5 5' XhoI site. The transgene can then be excised with NotI and
purified from vector sequences by gel electrophoresis. The
resulting transgene construct will contain the human J-Cκ
intronic enhancer and may contain the human 3' κ enhancer.

10 2. pCK2, a Cκ vector with heavy chain enhancers for
expressing rearranged variable segments
        A 0.9 kb XbaI fragment of mouse genomic DNA
containing the mouse heavy chain J-μ intronic enhancer (J.
Banerji et al., Cell 33:729-740 (1983)) was subcloned into
15 pUC18 to generate the plasmid pJH22.1. This plasmid was
linearized with SphI and the ends filled in with Klenow
enzyme. The Klenow treated DNA was then digested with HindIII
and a 1.4 kb MluI/HindIII fragment of phage clone λ1.3
(previous example), containing the human heavy chain J-μ
20 intronic enhancer (Hayday et al., Nature 307:334-340 (1984)),
to it. The resulting plasmid, pMHE1, consists of the mouse
and human heavy chain J-μ intronic enhancers ligated together
into pUC18 such that they are excised on a single
BamHI/HindIII fragment. This 2.3 kb fragment is isolated and
25 cloned into pGP1c to generate pMHE2. pMHE2 is digested with
SalI and the 13 kb XhoI insert of p36.5 cloned in. The
resulting plasmid, pCK2, is identical to pCK1, except that the
mouse and human heavy chain J-μ intronic enhancers are fused
to the 3' end of the transgene insert. To modulate expression
30 of the final transgene, analogous constructs can be generated
with different enhancers, i.e. the mouse or rat 3' kappa or
heavy chain enhancer (Meyer and Neuberger, EMBO J.,
8:1959-1964 (1989); Petterson et al., Nature, 344:165-168
(1990)).

35

112

3. Isolation of rearranged kappa light chain variable segments

        Two human leukocyte genomic DNA libraries cloned into the phage vector λEMBL3/SP6/T7 (Clonetech Laboratories, Inc., Palo Alto, CA) were screened with the human kappa light chain J region containing 3.5 kb XhoI/SmaI fragment of p36.5. Positive clones were tested for hybridization with the following Vκ specific oligonucleotide:

oligo-65 5'-agg ttc agt ggc agt ggg tct ggg aca gac ttc act
          ctc acc atc agc-3'

Clones that hybridized with both V and J probes are isolated and the DNA sequence of the rearranged VJκ segment determined.

4. Generation of transgenic mice containing rearranged human light chain constructs.

        Fragments containing functional VJ segments (open reading frame and splice signals) are subcloned into the unique XhoI sites of vectors pCK1 and pCK2 to generate rearranged kappa light chain transgenes. The transgene constructs are isolated from vector sequences by digestion with NotI. Agarose gel purified insert is microinjected into mouse embryo pronuclei to generate transgenic animals. Animals expressing human kappa chain are bred with heavy chain minilocus containing transgenic animals to generate mice expressing fully human antibodies.

        Because not all VJκ combinations may be capable of forming stable heavy-light chain complexes with a broad spectrum of different heavy chain VDJ combinations, several different light chain transgene constructs are generated, each using a different rearranged VJk clone, and transgenic mice that result from these constructs are bred with heavy chain minilocus transgene expressing mice. Peripheral blood, spleen, and lymph node lymphocytes are isolated from double transgenic (both heavy and light chain constructs) animals, stained with fluorescent antibodies specific for human and mouse heavy and light chain immunoglobulins (Pharmingen, San

113

Diego, CA) and analyzed by flow cytometry using a FACScan analyzer (Becton Dickinson, San Jose, CA). Rearranged light chain transgenes constructs that result in the highest level of human heavy/light chain complexes on the surface of the

5   highest number of B cells, and do not adversely affect the immune cell compartment (as assayed by flow cytometric analysis with B and T cell subset specific antibodies), are selected for the generation of human monoclonal antibodies.

10 D. Construction of unrearranged light chain minilocus
    transgenes
        1. pJCK1, a Jκ, Cκ containing vector for constructing
    minilocus transgenes
            The 13 kb Cκ containing XhoI insert of p36.5 is

15 treated with Klenow enzyme and cloned into HindIII digested, Klenow-treated, plasmid pGP1d. A plasmid clone is selected such that the 5' end of the insert is adjacent to the vector derived ClaI site. The resulting plasmid, p36.5-1d, is digested with ClaI and Klenow-treated. The Jκ1 containing 7.4

20 kb XhoI insert of p36.2 is then Klenow-treated and cloned into the ClaI, Klenow-treated p36.5-1d. A clone is selected in which the p36.2 insert is in the same orientation as the p36.5 insert. This clone, pJCK1 (Fig. 34), contains the entire human Jκ region and Cκ, together with 7.2 kb of upstream

25 sequences and 9 kb of downstream sequences. The insert also contains the human J-Cκ intronic enhancer and may contain a human 3' κ enhancer. The insert is flanked by a unique 3' SalI site for the purpose of cloning additional 3' flanking sequences such as heavy chain or light chain enhancers. A

30 unique XhoI site is located at the 5' end of the insert for the purpose of cloning in unrearranged Vκ gene segments. The unique SalI and XhoI sites are in turn flanked by NotI sites that are used to isolate the completed transgene construct away from vector sequences.

35

114

2. Isolation of unrearranged Vκ gene segments and generation
of transgenic animals expressing human Ig light chain protein

        The Vκ specific oligonucleotide, oligo-65 (discussed
above), is used to probe a human placental genomic DNA library
5   cloned into the phage vector lEMBL3/SP6/T7 (Clonetech
Laboratories, Inc., Palo Alto, CA). Variable gene segments
from the resulting clones are sequenced, and clones that
appear functional are selected. Criteria for judging
functionality include: open reading frames, intact splice
10 acceptor and donor sequences, and intact recombination
sequence. DNA fragments containing selected variable gene
segments are cloned into the unique XhoI site of plasmid pJCK1
to generate minilocus constructs. The resulting clones are
digested with NotI and the inserts isolated and injected into
15 mouse embryo pronuclei to generate transgenic animals. The
transgenes of these animals will undergo V to J joining in
developing B-cells. Animals expressing human kappa chain are
bred with heavy chain minilocus containing transgenic animals
to generate mice expressing fully human antibodies.

20

                        EXAMPLE 15

Genomic Heavy Chain Human Ig Transgene
        This Example describes the cloning of a human
genomic heavy chain immunoglobulin transgene which is then
25 introduced into the murine germline via microinjection into
zygotes or integration in ES cells.
        Nuclei are isolated from fresh human placental
tissue as described by Marzluff, W.F., et al. (1985),
Transcription and Translation: A Practical Approach, B.D.
30 Hammes and S.J. Higgins, eds., pp. 89-129, IRL Press, Oxford).
     The isolated nuclei (or PBS washed human spermatocytes) are
embedded in 0.5% low melting point agarose blocks and lyse
with 1 mg/ml proteinase K in 500mM EDTA, 1% SDS for nuclei, or
with 1mg/ml proteinase K in 500mM EDTA, 1% SDS, 10mM DTT for
35 spermatocytes at 50°C for 18 hours. The proteinase K is
inactivated by incubating the blocks in 40μg/ml PMSF in TE for
30 minutes at 50°C, and then washing extensively with TE. The
DNA is then digested in the agarose with the restriction

115

enzyme NotI as described by M. Finney in Current Protocols in Molecular Biology (F. Ausubel et al., eds. John Wiley & Sons, Supp. 4, 1988, e.g., Section 2.5.1).

     The NotI digested DNA is then fractionated by pulsed
5 field gel electrophoresis as described by Anand et al., Nuc. Acids Res. 17:3425-3433 (1989). Fractions enriched for the NotI fragment are assayed by Southern hybridization to detect one or more of the sequences encoded by this fragment. Such sequences include the heavy chain D segments, J segments, and
10 γ1 constant regions together with representatives of all 6 V_H families (although this fragment is identified as 670 kb fragment from HeLa cells by Berman et al. (1988), supra., we have found it to be an 830 kb fragment from human placental and sperm DNA). Those fractions containing this NotI
15 fragment are ligated into the NotI cloning site of the vector pYACNN as described (McCormick et al., Technique 2:65-71 (1990)). Plasmid pYACNN is prepared by digestion of pYACneo (Clontech) with EcoRI and ligation in the presence of the oligonucleotide 5' - AAT TGC GGC CGC - 3'.

20   YAC clones containing the heavy chain NotI fragment are isolated as described by Traver et al., Proc. Natl. Acad. Sci. USA, 86:5898-5902 (1989). The cloned NotI insert is isolated from high molecular weight yeast DNA by pulse field gel electrophoresis as described by M. Finney, op. cit. The
25 DNA is condensed by the addition of 1 mM spermine and microinjected directly into the nucleus of single cell embryos previously described. Alternatively, the DNA is isolated by pulsed field gel electrophoresis and introduced into ES cells by lipofection (Gnirke et al., EMBO J. 10:1629-1634 (1991)),
30 or the YAC is introduced into ES cells by spheroplast fusion.

EXAMPLE 16

Discontinuous Genomic Heavy Chain Ig Transgene

     An 85 kb SpeI fragment of human genomic DNA,
35 containing V_H6, D segments, J segments, the μ constant region and part of the γ constant region, has been isolated by YAC cloning essentially as described in Example 1. A YAC carrying a fragment from the germline variable region, such as a 570 kb

116

NotI fragment upstream of the 6 -830 kb NotI fragment described above containing multiple copies of V_1 through V_5, is isolated as described. (Berman et al. (1988), supra. detected two 570 kb NotI fragments, each containing multiple V

5 segments.) The two fragments are coinjected into the nucleus of a mouse single cell embryo as described in Example 1.

Typically, coinjection of two different DNA fragments result in the integration of both fragments at the same insertion site within the chromosome. Therefore,

10 approximately 50% of the resulting transgenic animals that contain at least one copy of each of the two fragments will have the V segment fragment inserted upstream of the constant region containing fragment. Of these animals, about 50% will carry out V to DJ joining by DNA inversion and about 50% by

15 deletion, depending on the orientation of the 570 kb NotI fragment relative to the position of the 85 kb SpeI fragment. DNA is isolated from resultant transgenic animals and those animals found to be containing both transgenes by Southern blot hybridization (specifically, those animals containing

20 both multiple human V segments and human constant region genes) are tested for their ability to express human immunoglobulin molecules in accordance with standard techniques.

25   EXAMPLE 17

Identification of functionally rearranged variable region sequences in transgenic B cells

An antigen of interest is used to immunize (see Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring

30 Harbor, New York (1988)) a mouse with the following genetic traits: homozygosity at the endogenous having chain locus for a deletion of J_H (Examples 10); hemizygous for a single copy of unrearranged human heavy chain minilocus transgene (examples 5 and 14); and hemizygous for a single copy of a

35 rearranged human kappa light chain transgene (Examples 6 and 14).

Following the schedule of immunization, the spleen is removed, and spleen cells used to generate hybridomas.

117

Cells from an individual hybridoma clone that secretes
antibodies reactive with the antigen of interest are used to
prepare genomic DNA. A sample of the genomic DNA is digested
with several different restriction enzymes that recognize
5   unique six base pair sequences, and fractionated on an agarose
gel. Southern blot hybridization is used to identify two DNA
fragments in the 2-10 kb range, one of which contains the
single copy of the rearranged human heavy chain VDJ sequences
and one of which contains the single copy of the rearranged
10   human light chain VJ sequence. These two fragments are size
fractionated on agarose gel and cloned directly into pUC18.
The cloned inserts are then subcloned respectively into heavy
and light chain expression cassettes that contain constant
region sequences.

15   The plasmid clone pγe1 (Example 12) is used as a
heavy chain expression cassette and rearranged VDJ sequences
are cloned into the XhoI site. The plasmid clone pCK1 is used
as a light chain expression cassette and rearranged VJ
sequences are cloned into the XhoI site. The resulting clones
20   are used together to transfect SP$_0$ cells to produce antibodies
that react with the antigen of interest (Co. et al., _Proc._
Natl. Acad. Sci. USA __88__:2869 (1991), which is incorporated
herein by reference).

          Alternatively, mRNA is isolated from the cloned
25   hybridoma cells described above, and used to synthesize cDNA.
The expressed human heavy and light chain VDJ and VJ sequence
are then amplified by PCR and cloned (Larrick et al., _Biol._
Technology, __7__:934-938 (1989)). After the nucleotide sequence
of these clones has been determined, oligonucleotides are
30   synthesized that encode the same polypeptides, and synthetic
expression vectors generated as described by Queen et al.,
Proc. Natl. Acad. Sci. USA., __84__:5454-5458 (1989).

Immunization of Transgenic Animals with Complex Antigens

35   The following experiment demonstrates that
transgenic animals can be successfully immunized with complex
antigens such as those on human red blood cells and respond

118

with kinetics that are similar to the response kinetics observed in normal mice.

Blood cells generally are suitable immunogens and comprise many different types of antigens on the surface of

5   red and white blood cells.

Immunization with human blood

Tubes of human blood from a single donor were collected and used to immunize transgenic mice having

10 functionally disrupted endogenous heavy chain loci (J_HD) and harboring a human heavy chain minigene construct (HC1); these mice are designated as Line 112. Blood was washed and resuspended in 50 mls Hanks' and diluted to $1x10^8$ cells/ml 0.2 mls ($2x10^7$ cells) were then injected interperitoneally using a

15 28 gauge needle and 1 cc syringe. This immunization protocol was repeated approximately weekly for 6 weeks. Serum titers were monitored by taking blood from retro-orbital bleeds and collecting serum and later testing for specific antibody. A pre-immune bleed was also taken as a control. On the very

20 last immunization, three days before these animals were sacrificed for serum and for hybridomas, a single immunization of 1×10^7 cells was given intravenously through the tail to enhance the production of hybridomas.

25

Table 9

Animals

	Mouse ID	Line	Sex	HC1-112	JHD
1	2343	112	M	+	++
2	2344	112	M	-	+
3	2345	112	F	-	+
4	2346	112	F	-	++
5	2347	112	F	-	++
6	2348	112	F	+	++
7	2349	112	F	-	+

30

35

Mice # 2343 and 2348 have a desired phenotype: human heavy
chain mini-gene transgenic on heavy chain knock-out
background.

5   Generation of Hybridomas

        Hybridomas were generated by fusing mouse spleen
cells of approximately 16 week-old transgenic mice (Table 9)
that had been immunized as described (supra) to a fusion
partner consisting of the non-secreting HAT-sensitive myeloma
10   cell line, X63 Ag8.653. Hybridoma clones were cultivated and
hybridoma supernatants containing immunoglobulins having
specific binding affinity for blood cell antigens were
identified, for example, by flow cytometry.

15   Flow cytometry

        Serum and hybridoma supernatants were tested using
flow cytometry. Red blood cells from the donor were washed 4X
in Hanks' balanced salt solution and 50,000 cells were placed
in 1.1 ml polypropylene microtubes. Cells were incubated with
20   antisera or supernatant from the hybridomas for 30 minutes on
ice in staining media (1x RPMI 1640 media without phenol red
or biotin (Irvine Scientific) 3% newborn calf serum, 0.1% Na
azide). Controls consisted of littermate mice with other
genotypes. Cells were then washed by centrifugation at 4°C in
25   Sorvall RT600B for 5-10 minutes at 1000 rpm. Cells were
washed two times and then antibody detected on the cell
surface with a fluorescent developing reagent. Two monoclonal
reagents were used to test. One was a FITC-labeled mouse
anti-human μ heavy chain antibody (Pharmagen, San Diego, CA)
30   and the other was a PE-labeled rat anti-mouse kappa light
chain (Becton-Dickenson, San Jose, CA). Both of these
reagents gave similar results. Whole blood (red blood cells
and white blood cells) and white blood cells alone were used
as target cells. Both sets gave positive results.
35   Serum of transgenic mice and littermate controls was
incubated with either red blood cells from the donor, or white
blood cells from another individual, washed and then developed
with anti-human IgM FITC labeled antibody and analyzed in a

flow cytometer. Results showed that serum from mice that are transgenic for the human mini-gene locus (mice 2343 and 2348) show human IgM reactivity whereas all littermate animals (2344, 2345, 2346, 2347) do not. Normal mouse serum (NS) and

5   phosphate buffer saline (PBS) were used as negative controls. Red blood cells were ungated and white blood cells were gated to include only lymphocytes. Lines are drawn on the x and y axis to provide a reference. Flow cytometry was performed on 100 supernatants from fusion 2348. Four supernatants showed

10 positive reactivity for blood cell antigens.

EXAMPLE 18

Reduction of Endogenous Mouse Immunoglobulin Expression
by Antisense RNA

15   A.   Vector for Expression of Antisense Ig Sequences
            1.   Construction of the cloning vector pGP1h
            The vector pGP1b (referred to in a previous example) is digested with XhoI and BamHI and ligated with the following oligonucleotides:

20

    5'- gat cct cga gac cag gta cca gat ctt gtg aat tcg -3'
    5'- tcg acg aat tca caa gat ctg gta cct ggt ctc gag -3'

    to generate the plasmid pGP1h. This plasmid contains a

25 polylinker that includes the following restriction sites: NotI, EcoRI, BglII, Asp718, XhoI, BamHI, HindIII, NotI.

        Construction of pBCE1.
        A 0.8 kb XbaI/BglII fragment of pVH251 (referred to

30 in a previous example), that includes the promoter leader sequence exon, first intron, and part of the second exon of the human VH-V family immunoglobulin variable gene segment, was inserted into XbaI/BglII digested vector pNN03 to generate the plasmid pVH251.

35   The 2.2 kb BamHI/EcoRI DNA fragment that includes the coding exons of the human growth hormone gene (hGH; Seeburg, (1982) DNA 1:239-249) is cloned into BglII/EcoRI digested pGH1h. The resulting plasmid is digested with BamHI

121

and the BamHI/BglII of pVH251N is inserted in the same
orientation as the hGH gene to generate the plasmid pVhgh.

A 0.9 kb XbaI fragment of mouse genomic DNA
containing the mouse heavy chain J-μ intronic enhancer
5   (Banerji et al., (1983) Cell 33:729-740) was subcloned into
pUC18 to generate the plasmid pJH22.1. This plasmid was
linearized with SphI and the ends filled in with klenow
enzyme. The klenow treated DNA was then digested with HindIII
and a 1.4 kb MluI(klenow)/HindIII fragment of phage clone λ1.3
10 (previous example), containing the human heavy chain J-μ
intronic enhancer (Hayday et al., (1984) Nature 307:334-340),
to it.   The resulting plasmid, pMHE1, consists of the mouse
and human heavy chain J-μ intron enhancers ligated together
into pUC18 such that they can be excised on a single
15 BamHI/HindIII fragment.

The BamHI/HindIII fragment of pMHE1 is cloned into
BamHI/HindIII cut pVhgh to generate the B-cell expression
vector pBCE1. This vector, depicted in Fig. 36, contains
unique XhoI and Asp718 cloning sites into which antisense DNA
20 fragments can be cloned. The expression of these antisense
sequences is driven by the upstream heavy chain promoter-
enhancer combination the downstream hGH gene sequences provide
polyadenylation sequences in addition to intron sequences that
promote the expression of transgene constructs. Antisense
25 transgene constructs generated from pBCE1 can be separated
from vector sequences by digestion with NotI.

B.   An IgM antisense transgene construct.
The following two oligonucleotides:

30
5'- cgc ggt acc gag agt cag tcc ttc cca aat gtc -3'
5'- cgc ctc gag aca gct gga atg ggc aca tgc aga -3'

are used as primers for the amplification of mouse IgM
35 constant region sequences by polymerase chain reaction (PCR)
using mouse spleen cDNA as a substrate. The resulting 0.3 kb
PCR product is digested with Asp718 and XhoI and cloned into
Asp718/XhoI digested pBCE1 to generate the antisense transgene

122

construct pMAS1. The purified NotI insert of pMAS1 is
microinjected into the pronuclei of half day mouse embryos--
alone or in combination with one or more other transgene
constructs--to generate transgenic mice. This construct

5   expresses an RNA transcript in B-cells that hybridizes with
mouse IgM mRNA, thus down-regulating the expression of mouse
IgM protein. Double transgenic mice containing pMAS1 and a
human heavy chain transgene minilocus such as pHC1 (generated
either by coinjection of both constructs or by breeding of

10 singly transgenic mice) will express the human transgene
encoded Ig receptor on a higher percentage of B-cell than mice
transgenic for the human heavy chain minilocus alone. The
ratio of human to mouse Ig receptor expressing cells is due in
part to competition between the two populations for factors

15 and cells that promoter B-cell differentiation and expansion.
Because the Ig receptor plays a key role in B-cell
development, mouse Ig receptor expressing B-cells that express
reduced levels of IgM on their surface (due to mouse Ig
specific antisense down-regulation) during B-cell development

20 will not compete as well as cells that express the human
receptor.

    C.   An IgKappa antisense transgene construct.
         The following two oligonucleotides:

25
    5'- cgc ggt acc gct gat gct gca cca act gta tcc -3'
    5'- cgc ctc gag cta aca ctc att cct gtt gaa gct -3'

    are used as primers for the amplification of mouse IgKappa
30 constant region sequences by polymerase chain reaction (PCR)
    using mouse spleen cDNA as a substrate. The resulting 0.3 kb
    PCR product is digested with Asp718 and XhoI and cloned into
    Asp718/XhoI digested pBCE1 to generate the antisense transgene
    construct pKAS1. The purified NotI insert of pKAS1 is
35 microinjected into the pronuclei of half day mouse embryos--
    alone or in combination with one or more other transgene
    constructs--to generate transgenic mice. This construct
    expresses an RNA transcript in B-cells that hybridizes with

123

mouse IgK mRNA, thus down-regulating the expression of mouse
IgK protein as described above for pMAS1.

EXAMPLE 19

5   This example demonstrates the successful
immunization and immune response in a transgenic mouse of the
present invention.

Immunization of Mice

10   Keyhole limpet hemocyanin conjugated with greater
than 400 dinitrophenyl groups per molecule (Calbiochem, La
Jolla, California) (KLH-DNP) was alum precipitated according
to a previously published method (Practical Immunology, L.
Hudson and F.C. Hay, Blackwell Scientific (Pubs.), p. 9,
15 1980). Four hundred µg of alum precipitated KLH-DNP along
with 100 µg dimethyldioctadecyl Ammonium Bromide in 100 µL of
phosphate buffered saline (PBS) was injected intraperitoneally
into each mouse. Serum samples were collected six days later
by retro-orbital sinus bleeding.

20

Analysis of Human Antibody Reactivity in Serum

          Antibody reactivity and specificity were assessed
using an indirect enzyme-linked immunosorbent assay (ELISA).
Several target antigens were tested to analyze antibody
25 induction by the immunogen. Keyhole limpet hemocyanin
(Calbiochem) was used to identify reactivity against the
protein component, bovine serum albumin-DNP for reactivity
against the hapten and/or modified amino groups, and KLH-DNP
for reactivity against the total immunogen. Human antibody
30 binding to antigen was detected by enzyme conjugates specific
for IgM and IgG sub-classes with no cross reactivity to mouse
immunoglobulin. Briefly, PVC microtiter plates were coated
with antigen drying overnight at 37°C of 5 µg/mL protein in
PBS. Serum samples diluted in PBS, 5% chicken serum, 0.5%
35 Tween-20 were incubated in the wells for 1 hour at room
temperature, followed by anti-human IgG Fc and IgG F(ab')-
horseradish peroxidase or anti-human IgM Fc-horseradish
peroxidase in the same diluent. After 1 hour at room

124

temperature enzyme activity was assessed by addition of ABTS substrate (Sigma, St. Louis, Missouri) and read after 30 minutes at 415-490 nm.

5   Human Heavy Chain Participation in Immune Response in Transgenic Mice

Figure 37 illustrates the response of three mouse littermates to immunization with KLH-DNP. Mouse number 1296 carried the human IgM and IgG unrearranged transgene and was
10 homozygous for mouse Ig heavy chain knockout. Mouse number 1299 carried the transgene on a non-knockout background, while mouse 1301 inherited neither of these sets of genes. Mouse 1297, another littermate, carried the human transgene and was hemizygous with respect to mouse heavy chain knockout. It was
15 included as a non-immunized control.

The results demonstrate that both human IgG and IgM responses were developed to the hapten in the context of conjugation to protein. Human IgM also developed to the KLH molecule, but no significant levels of human IgG were present
20 at this time point. In pre-immunization serum samples from the same mice, titers of human antibodies to the same target antigens were insignificant.

EXAMPLE 20

This example demonstrates the successful
25 immunization with a human antigen and immune response in a transgenic mouse of the present invention, and provides data demonstrating that nonrandom somatic mutation occurs in the variable region sequences of the human transgene.

30
Demonstration of antibody responses comprising human immunoglobulin heavy chains against a human glycoprotein antigen

Transgenic mice used for the experiment were
35 homozygous for functionally disrupted murine immunoglobulin heavy chain loci produced by introduction of a transgene at the joining (J) region (supra) resulting in the absence of functional endogenous (murine) heavy chain production. The

125

transgenic mice also harbored at least one complete
unrearranged human heavy chain mini-locus transgene, (HC1,
supra), which included a single functional V_H gene (V_H251),
human μ constant region gene, and human γ1 constant region
5   gene. Transgenic mice shown to express human immunoglobulin
transgene products (supra) were selected for immunization with
a human antigen to demonstrate the capacity of the transgenic
mice to make an immune response against a human antigen
immunization. Three mice of the HC1-26 line and three mice of
10   the HC1-57 line (supra) were injected with human antigen.

One hundred μg of purified human carcinoembryonic
antigen (CEA) insolubilized on alum was injected in complete
Freund's adjuvant on Day 0, followed by further weekly
injections of alum-precipitated CEA in incomplete Freund's
15   adjuvant on Days 7, 14, 21, and 28. Serum samples were
collected by retro-orbital bleeding on each day prior to
injection of CEA. Equal volumes of serum were pooled from
each of the three mice in each group for analysis.

Titres of human μ chain-containing immunoglobulin
20   and human γ chain-containing immunoglobulin which bound to
human CEA immobilized on microtitre wells were determined by
ELISA assay. Results of the ELISA assays for human μ chain-
containing immunoglobulins and human γ chain-containing
immmunoglbulins are shown in Figs. 38 and 39, respectively.
25   Significant human μ chain Ig titres were detected for both
lines by Day 7 and were observed to rise until about Day 21.
For human γ chain Ig, significant titres were delayed, being
evident first for line HC1-57 at Day 14, and later for line
HC1-26 at Day 21. Titres for human γ chain Ig continued to
30   show an increase over time during the course of the
experiment. The observed human μ chain Ig response, followed
by a plateau, combined with a later geveloping γ chain
response which continues to rise is characteristic of the
pattern seen with affinity maturation. Analysis of Day 21
35   samples showed lack of reactivity to an unrelated antigen,
keyhole limpet hemocyanin (KLC), indicating that the antibody
response was directed against CEA in a specific manner.

126

These data indicate that animals transgenic for human unrearranged immunoglobulin gene loci: (1) can respond to a human antigen (e.g., the human glycoprotein, CEA), (2) can undergo isotype switching ("class switching) as

5 exemplified by the observed μ to γ class switch, and (3) exhibit characteristics of affinity maturation in their humoral immune responses. In general, these data indicate: (1) the human Ig transgenic mice have the ability to induce heterologous antibody production in response to a defined

10 antigen, (2) the capacity of a single transgene heavy chain variable region to respond to a defined antigen, (3) response kinetics over a time period typical of primary and secondary response development, (4) class switching of a transgene-encoded humoral immune response from IgM to IgG, and (5) the

15 capacity of transgenic animal to produce human-sequence antibodies against a human antigen.

Demonstration of somatic mutation in a human heavy chain transgene minilocus.

20   Line HC1-57 transgenic mice, containing multiple copies of the HC1 transgene, were bred with immunoglobulin heavy chain deletion mice to obtain mice that contain the HC1 transgene and contain disruptions at both alleles of the endogenous mouse heavy chain (supra). These mice express

25 human mu and gamma1 heavy chains together with mouse kappa and lambda light chains (supra). One of these mice was hyperimmunized against human carcinoembryonic antigen by repeated intraperitoneal injections over the course of 1.5 months. This mouse was sacrificed and lymphoid cells isolated

30 from the spleen, inguinal and mesenteric lymph nodes, and peyers patches. The cells were combined and total RNA isolated. First strand cDNA was synthesized from the RNA and used as a template for PCR amplification with the following 2 oligonucleotide primers:

35

149   5'-cta gct cga gtc caa gga gtc tgt gcc gag gtg cag ctg (g/a/t/c)-3'

151   5'-ggc gct cga gtt cca cga cac cgt cac cgg ttc-3'

127

These primers specifically amplify VH251/gamma1 cDNA
sequences. The amplified sequences were digested with XhoI
and cloned into the vector pNN03. DNA sequence from the
inserts of 23 random clones is shown in Fig. 40; sequence
5 variations from germline sequence are indicated, dots indicate
sequence is identical to germline. Comparison of the cDNA
sequences with the germline sequence of the VH251 transgene
reveals that 3 of the clones are completely unmutated, while
the other 20 clones contain somatic mutations. One of the 3
10 non-mutated sequences is derived from an out-of-frame VDJ
joint. Observed somatic mutations at specific positions of
occur at similar frequencies and in similar distribution
patterns to those observed in human lymphocytes (Cai et al.
(1992) J. Exp. Med. 176: 1073, incorporated herein by
15 reference). The overall frequency of somatic mutations is
approximately 1%; however, the frequency goes up to about 5%
within CDR1, indicating selection for amino acid changes that
affect antigen binding. This demonstrates antigen driven
affinity maturation of the human heavy chain sequences.

20

EXAMPLE 21

This example demonstrates the successful formation
of a transgene by co-introduction of two separate
polynucleotides which recombine to form a complete human light
25 chain minilocus transgene.

Generation of an unrearranged light chain minilocus transgene
by co-injection of two overlapping DNA fragments
1.   Isolation of unrearranged functional V_κ gene segments
30   vk65.3, vk65.5, vk65.8 and vk65.15
The V_κ specific oligonucleotide, oligo-65 (5'-agg
ttc agt ggc agt ggg tct ggg aca gac ttc act ctc acc atc agc-
3'), was used to probe a human placental genomic DNA library
cloned into the phage vector λEMBL3/SP6/T7 (Clonetech
35 Laboratories, Inc., Palo Alto, CA). DNA fragments containing
V_κ segments from positive phage clones were subcloned into
plasmid vectors. Variable gene segments from the resulting
clones are sequenced, and clones that appear functional were

128

selected. Criteria for judging functionality include: open
reading frames, intact splice acceptor and donor sequences,
and intact recombination sequence. DNA sequences of 4
functional V_κ gene segments (vk65.3, vk65.5, vk65.8, and
5   vk65.15) from 4 different plasmid clones isolated by this
procedure are shown in Figs. 41-44. The four plasmid clones,
p65.3f, p65.5g1, p65.8, and p65.15f, are described below.

(1 a) p65.3f
10   A 3 kb Xba fragment of phage clone λ65.3 was
subcloned into pUC19 so that the vector derived SalI site was
proximal to the 3' end of the insert and the vector derived
BamHI site 5'. The 3 kb BamHI/SalI insert of this clone was
subcloned into pGP1f to generate p65.3f.

15

(1 b) p65.5g1
        A 6.8 kb EcoRI fragment of phage clone λ65.5 was
subcloned into pGP1f so that the vector derived XhoI site is
proximal to the 5' end of the insert and the vector derived
20 SalI site 3'. The resulting plasmid is designated p65.5g1.

(1 c) p65.8
        A 6.5 kb HindIII fragment of phage clone λ65.8 was
cloned into pSP72 to generate p65.8.

25

(1 d) p65.15f
        A 10 kb EcoRI fragment of phage clone λ65.16 was
subcloned into pUC18 to generate the plasmid p65.15.3. The V_κ
gene segment within the plasmid insert was mapped to a 4.6 kb
30 EcoRI/HindIII subfragment, which was cloned into pGP1f. The
resulting clone, p65.15f, has unique XhoI and SalI sites
located at the respective 5' and 3' ends of the insert.

2.   pKV4
35   The XhoI/SalI insert of p65.8 was cloned into the
XhoI site of p65.15f to generate the plasmid pKV2. The
XhoI/SalI insert of p65.5g1 was cloned into the XhoI site of
pKV2 to generate pKV3. The XhoI/SalI insert of pKV3 was

129

cloned into the XhoI site of p65.3f to generate the plasmid
pKV4. This plasmid contains a single 21 kb XhoI/SalI insert
that includes 4 functional V_κ gene segments. The entire
insert can also be excised with NotI.

5

3.   pKC1B

(3 a) pKcor

          Two XhoI fragments derived from human genomic DNA
phage λ clones were subcloned into plasmid vectors. The
10 first, a 13 kb $J_\kappa 2$-$J_\kappa 5$/C_κ containing fragment, was treated with
Klenow enzyme and cloned into HindIII digested, Klenow
treated, plasmid pGP1d. A plasmid clone (pK-31) was selected
such that the 5' end of the insert is adjacent to the vector
derived ClaI site. The second XhoI fragment, a 7.4 kb piece
15 of DNA containing $J_\kappa 1$ was cloned into XhoI/SalI-digested
pSP72, such that the 3' insert XhoI site was destroyed by
ligation to the vector SalI site. The resulting clone,
p36.2s, includes an insert derived ClaI site 4.5 kb upstream
of $J_\kappa 1$ and a polylinker derived ClaI site downstream in place
20 of the naturally occurring XhoI site between $J_\kappa 1$ and $J_\kappa 2$. This
clone was digested with ClaI to release a 4.7 kb fragment
which was cloned into ClaI digested pK-31 in the correct 5' to
3' orientation to generate a plasmid containing all 5 human J_κ
segments, the human intronic enhancer human C_κ, 4.5 kb of 5'
25 flanking sequence, and 9 kb of 3' flanking sequence. This
plasmid, pKcor, includes unique flanking XhoI and SalI sites
on the respective 5' and 3' sides of the insert.

(3 b) pKcorB

30   A 4 kb BamHI fragment containing the human 3' kappa
enhancer (Judde, J.-G. and Max, E.E. (1992) Mol. Cell. Biol.
12: 5206, incorporated herein by reference) was cloned into
pGP1f such that the 5' end is proximal to the vector XhoI
site. The resulting plasmid, p24Bf, was cut with XhoI and the
35 17.7 kb XhoI/SalI fragment of pKcor cloned into it in the same
orientation as the enhancer fragment. The resulting plasmid,
pKcorB, includes unique XhoI and SalI sites at the 5' and 3'
ends of the insert respectively.

130

(3 c) pKC1B

The XhoI/SalI insert of pKcorB was cloned into the SalI site of p65.3f to generate the light-chain minilocus-transgene plasmid pKC1B. This plasmid includes a single

5 functional human V_κ segment, all 5 human J_κ segments, the human intronic enhancer, human C_κ, and the human 3' kappa enhancer. The entire 25 kb insert can be isolated by NotI digestion.

4.   Co4

10 The two NotI inserts from plasmids pKV4 and pKC1B were mixed at a concentration of 2.5 μg/ml each in microinjection buffer, and co-injected into the pronuclei of half day mouse embryos as described in previous examples. Resulting transgenic animals contain transgene inserts

15 (designated Co4, product of the recombination shown in Fig. 45) in which the two fragments co-integrated. The 3' 3 kb of the pKV4 insert and the 5'3 kb of the pKC1B insert are identical. Some of the integration events will represent homologous recombinations between the two fragments over the 3

20 kb of shared sequence. The Co4 locus will direct the expression of a repertoire of human sequence light chains in a transgenic mouse.

The foregoing description of the preferred embodiments of the present invention has been presented for

25 purposes of illustration and description. They are not intended to be exhaustive or to limit the invention to the precise form disclosed, and many modifications and variations are possible in light of the above teaching.

All publications and patent applications herein are

30 incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the present invention has been described in some detail by way of illustration for purposes of clarity of

35 understanding, it will be apparent that certain changes and modifications may be practiced within the scope of the claims.

131

WHAT IS CLAIMED IS:

TRANSGENE CLAIMS

1.   An isolated immunoglobulin heavy chain transgene that is expressed in B cells of a transgenic nonhuman animal containing at least one integrated copy of a polynucleotide comprising a DNA sequence of the formula:

$$(V_H)_x-(D)_y-(J_H)_z-(S_D)_m-(C_1)_n-[(T)-(S_A)_p-(C_2)]_q$$

wherein x, y, z, m, n, p, and q are integers and x is 2-100, n is 1-10, y is 2-50, p is 1-10, z is 1-50, q is 0-50, and m is 0-10.

2.   A transgene of Claim 1, wherein said polynucleotide comprises at least one heterologous D gene segment that can be incorporated into a functionally rearranged V-D-J sequence.

3.   A transgene of Claim 2, wherein said heterologous D gene segment contains at least one human D gene.

4.   A transgene of Claim 1, wherein said polynucleotide comprises a human μ C_H gene segment and a human γ_1 C_H gene segment.

5.   A transgene of Claim 1, wherein q is at least 1, m is at least 1, n is at least 1, and said polynucleotide comprises at least about 50 basepairs of a segment immediately upstream of a germline switch sequence.

6.   A transgene of Claim 5, wherein said polynucleotide comprises about 200 basepairs of sequence immediately upstream of a human germline γ_1 switch sequence.

7.   A transgene of Claim 1, wherein a S_D segment is a γ_1 switch sequence.

132

8.   A transgene of Claim 1, wherein said polynucleotide comprises about 200 basepairs naturally upstream of a human germline γ_1 switch sequence, and wherein said 200 basepairs are operably linked to a human γ_1 switch
5 sequence.

9.   A transgene of Claim 1, wherein said D region comprises only heterologous D genes.

10 10.   A transgene of Claim 9, wherein said D region comprises only human D genes.

11.   A transgene of Claim 1, wherein the polynucleotide is functionally rearranged in vivo to produce a
15 rearranged V-D-J gene segment that contains a recognizable D region gene sequence.

12.   A transgenic non-human animal comprising a transgene of Claim 1 in the germline of said non-human animal.
20
13.   A transgenic non-human animal of Claim 12, wherein said transgene is rearranged.

14.   A transgenic non-human animal of Claim 12,
25 wherein said transgene is unrearranged.

15.   A transgenic non-human animal of Claim 14, wherein said B cells produce a heterologous antibody.

30 16.   A transgenic non-human animal of Claim 15, wherein said B cells produce a population of heterologous antibodies of more than one isotype.

17.   A transgenic non-human animal of Claim 12,
35 wherein said transgene encodes V_H, D, J_H, and C_H regions.

18.   A transgenic non-human animal of Claim 12 wherein said non-human animal is a rodent.

133

19.   A transgenic non-human animal of Claim 12, wherein serum of said animal comprises human antibodies that contain a recognizable D region gene sequence.

20.   A transgenic non-human animal of Claim 12, wherein at least one lymphocyte of said animal contains a mRNA encoding a heterologous immunoglobulin chain.

21.   A transgenic animal of claim 20, wherein said mRNA contains a recognizable D region gene sequence.

22.   A transgenic animal of claim 21, wherein said mRNA contains a functionally rearranged V-D-J sequence.

23.   A transgenic animal of claim 12, wherein said transgenic animal comprises heterologous antibodies which comprise a human-sequence μ chain and which specifically bind to an antigen.

24.   A transgenic animal of claim 23, further comprising heterologous antibodies which comprise a human-sequence γ chain and which specifically bind to an antigen.

25.   A transgenic animal of claim 24, wherein the antigen is a human antigen.

26.   A transgenic animal of claim 25, wherein the human antigen is CEA or a human blood cell antigen.

27.   A transgenic nonhuman animal comprising a human transgene of claim 1, wherein the transgenic animal further comprises lymphoid tissue containing a population of mRNA species having somatic mutations clustered in CDR regions of a variable region encoded by an immunoglobulin transgene.

28.   A transgenic nonhuman animal of claim 27, wherein the immunoglobulin transgene is a heavy chain minilocus corresponding to HC1.

134

29. A hybridoma comprising a transgenic non-human B cell fused with a second cell capable of immortalizing said B cell, wherein said hybridoma produces a monoclonal antibody heterologous to said non-human animal.

5

30. A hybridoma of Claim 29 wherein the heterologous antibody comprises a human heavy chain containing a recognizable D region gene sequence.

10   31. A hybridoma of Claim 29 wherein said B cell is of murine origin.

32. A hybridoma of Claim 29 wherein the monoclonal antibody binds to a human antigen with an affinity of at least 15 1×10^7 M^{-1}.

33. A human monoclonal antibody produced by a hybridoma of Claim 29.

34. A transgenic non-human animal having serum 20 comprising detectable heterologous antibodies and having at least one suppressed endogenous immunoglobulin locus.

35. A method for producing heterologous immunoglobulins from a transgenic nonhuman animal, the animal 25 having a genome comprising germline copies of at least one transgene of Claim 1, the method comprising:
        suppressing an endogenous immunoglobulin locus;
        contacting the animal with a preselected antigen; and
30   collecting said heterologous immunoglobulins.

36. A method according to Claim 35, wherein suppression is produced by an antisense polynucleotide.

35   37. A method according to Claim 36, wherein the antisense polynucleotide is transcribed from an integrated antisense transgene.

135

38. A method according to Claim 37, wherein transcription of the antisense transgene produces a transcript containing an antisense sequence linked to a second sequence.

39. A method according to Claim 37, wherein the non-human animal has a genome comprising germline copies of at least one light chain immunoglobulin transgene.

40. A method for suppressing at least one endogenous immunoglobulin locus in a transgenic non-human animal, comprising the steps of:

introducing an antisense transgene into a non-human animal to produce a non-human transgenic animal bearing an antisense transgene;

transcribing antisense RNA from the antisense transgene _in vivo_;

hybridizing the antisense RNA to a polynucleotide containing an endogenous immunoglobulin sequence; and

inhibiting expression of an endogenous immunoglobulin chain.

41. A method of Claim 40, wherein said antisense transgene contains a nucleotide sequence that is homologous to an endogenous kappa chain immunoglobulin gene sequence.

42. A method of Claim 40, wherein said antisense transgene contains a nucleotide sequence that is homologous to an endogenous heavy chain immunoglobulin gene sequence.

43. A method for inactivating an endogenous immunoglobulin gene, comprising the steps of:

integrating a targeting vector into an endogenous immunoglobulin gene; and

selecting for a cell bearing an integrated targeting vector.

44. A method according to Claim 43, further comprising the step of:

generating a line of non-human animals bearing a copy of the integrated targeting vector.

45.   A method according to Claim 43, wherein the endogenous immunoglobulin gene is a light chain gene.

46.   A method according to Claim 45, wherein the light chain gene is a kappa light chain gene.

47.   A method according to Claim 43, wherein the endogenous immunoglobulin gene is a heavy chain gene.

48.   A method according to Claim 43, wherein the endogenous immunoglobulin gene is a murine immunoglobulin gene.

49.   A method according to Claim 44, wherein the non-human animal is a mouse.

50.   A method according to Claim 49, wherein the mouse further comprises a human immunoglobulin transgene of Claim 1.

51.   A method according to Claim 43, wherein the cell is an embryonic stem cell.

52.   A method of generating a non-human animal having an inactivated endogenous immunoglobulin gene, comprising the steps of:
        breeding a line of non-human animals generated by the method of Claim 44; and
        identifying individual non-human animal offspring that are homozygous for an inactivated immunoglobulin gene.

53.   A method according to Claim 52, wherein said offspring have an inactivated heavy chain immunoglobulin gene and at least one inactivated light chain immunoglobulin gene.

137

54. A transgenic non-human animal with an
inactivated endogenous immunoglobulin gene.

55. A non-human animal of Claim 54, wherein an
integrated targeting vector is present in germline DNA.

56. A non-human animal of Claim 55, wherein a light
chain immunoglobulin gene is inactivated.

57. A non-human animal of Claim 55, wherein a heavy
chain immunoglobulin gene is inactivated.

58. A non-human animal of Claim 55 which is
homozygous for at least one inactivated endogenous
immunoglobulin gene.

59. An antisense transgene comprising a nucleotide
sequence that is complementary to a polynucleotide sequence
that is substantially identical to an immunoglobulin gene
sequence.

60. A transgenic non-human animal bearing an
antisense transgene of Claim 59.

61. A transgenic non-human animal comprising a
functionally disrupted endogenous heavy chain locus and a
heterologous immunoglobulin heavy chain transgene, wherein
said animal makes an antibody response following immunization
with an antigen.

62. A transgenic non-human animal of Claim 61,
wherein said functionally disrupted endogenous heavy chain
locus is a J_H region homologous recombination knockout, said
heterologous immunoglobulin heavy chain transgene is the HC1
human minigene transgene, and said antigen is a human antigen.

63. A transgenic non-human animal of Claim 61,
wherein the antibody response comprises a population of

antibodies which comprise human μ chain-containing immunoglobulins and human γ chain-containing immunoglobulins.

64. A transgenic non-human animal of Claim 63, wherein the heterologous antibodies comprise a population of heterologous immunoglobulins which comprise somatic mutation in the variable regions which cluster in the CDRs.

65. A transgenic non-human animal of Claim 63, wherein the antigen is selected from the group consisting of human blood cell surface antigens, KLH, and human CEA.
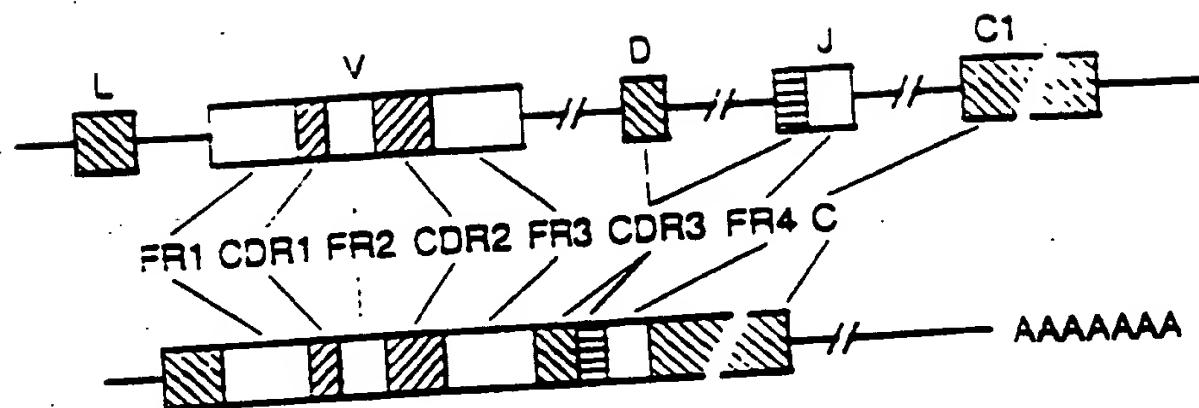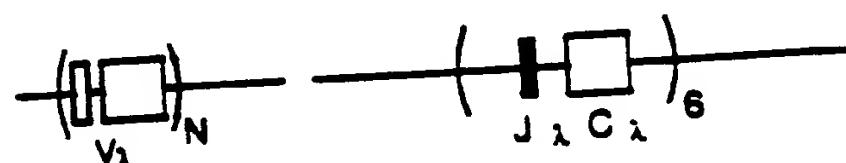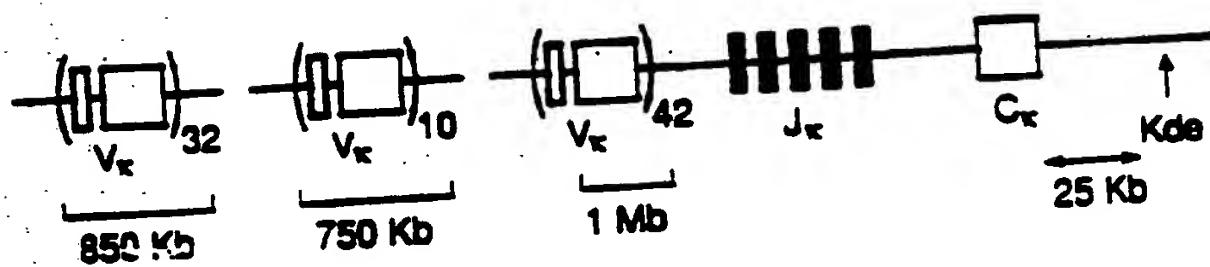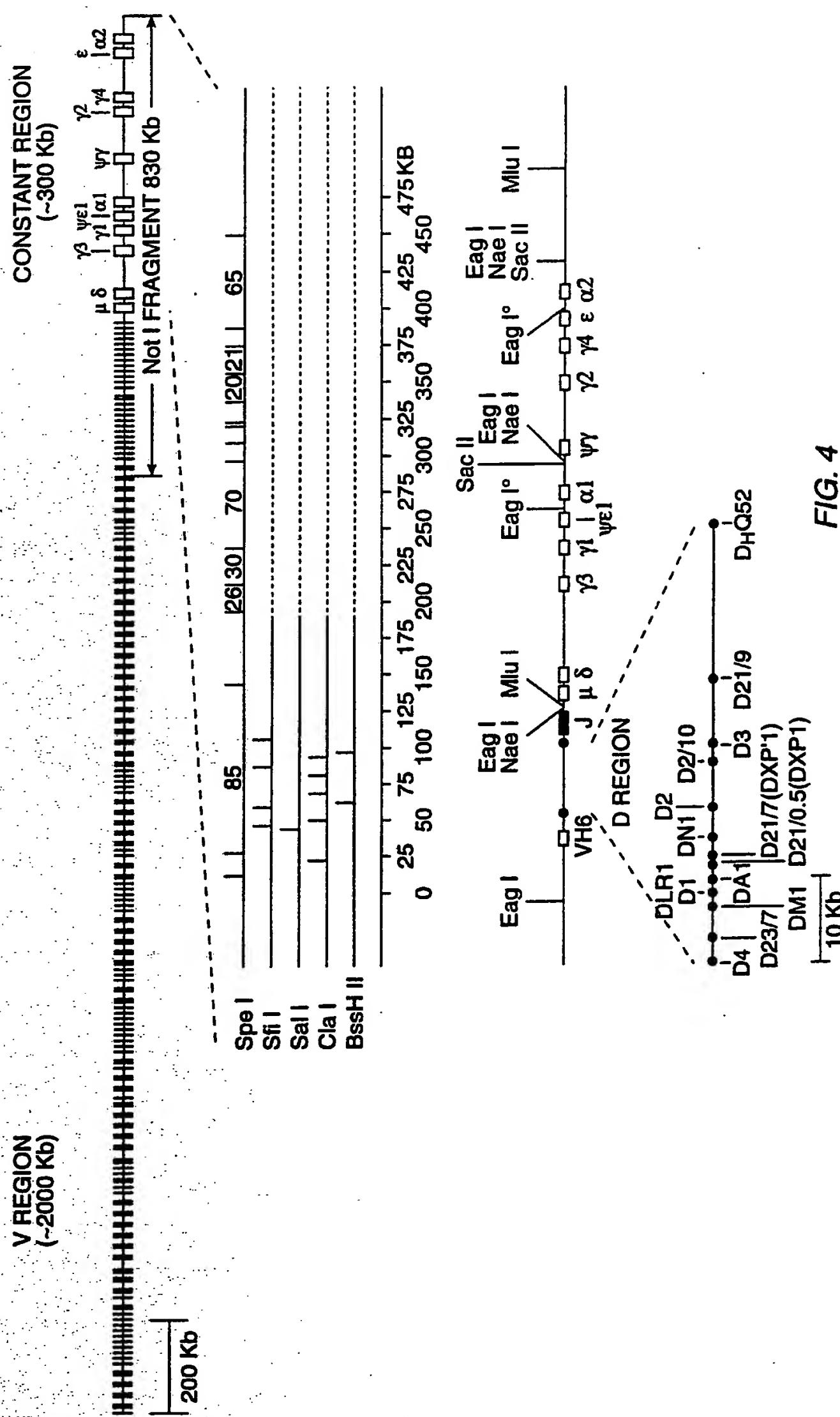
FIG. 1

FIG. 2

FIG. 3

FIG. 4
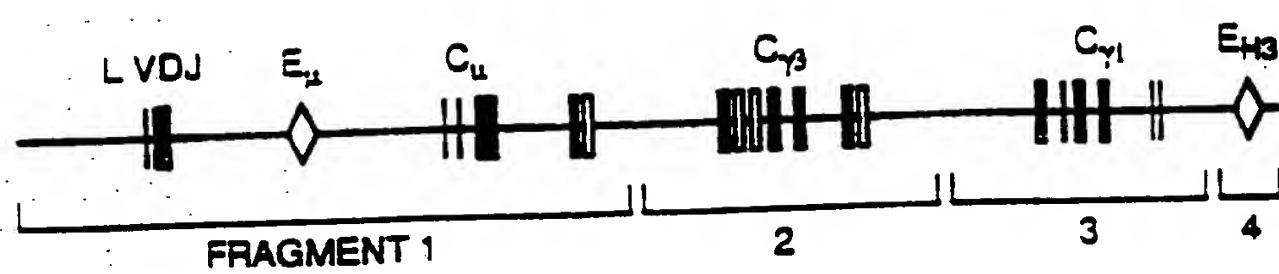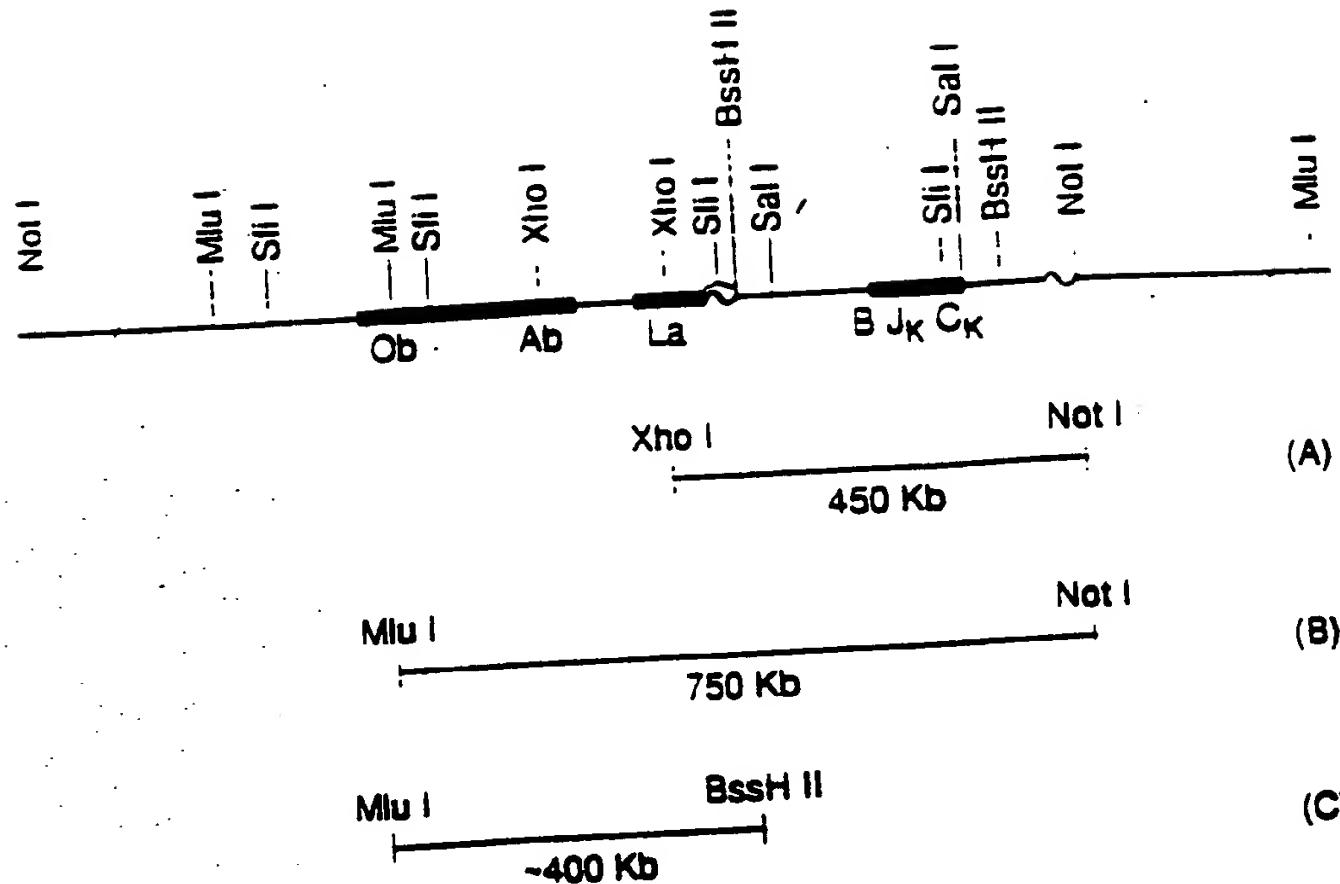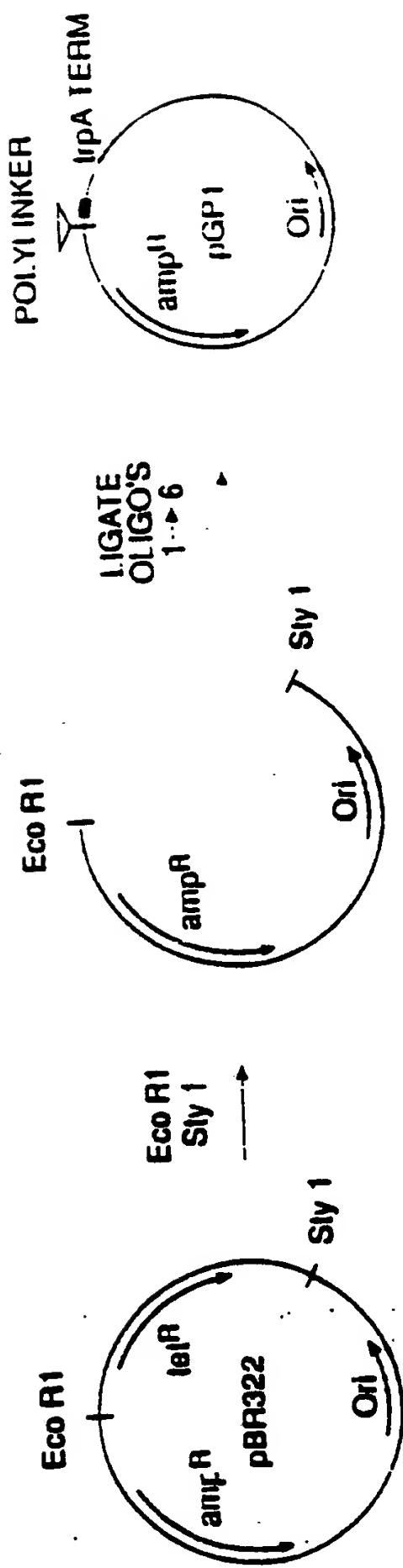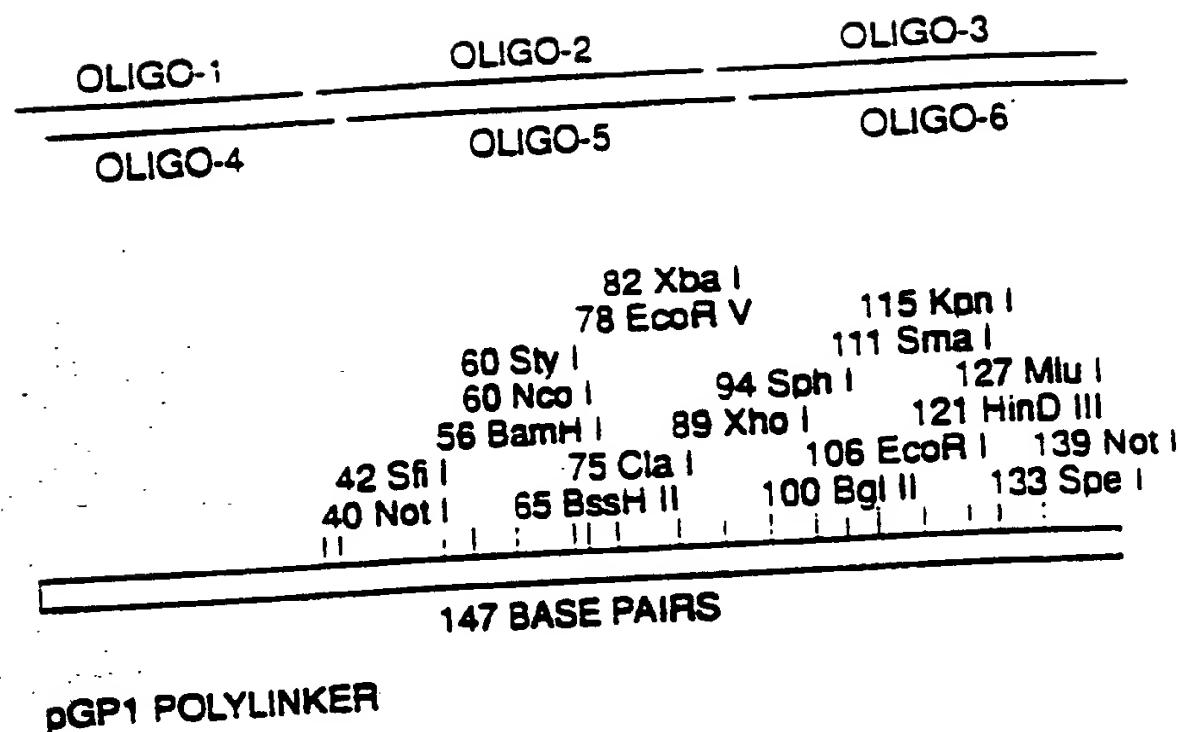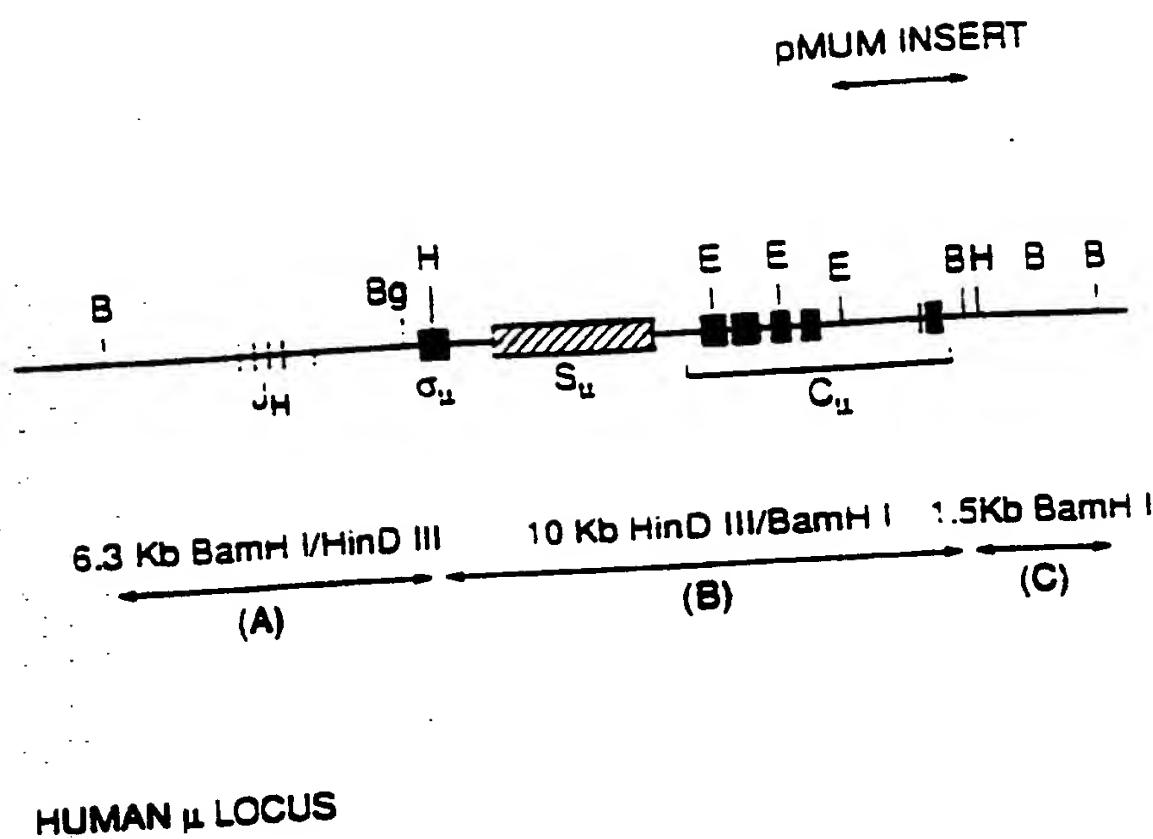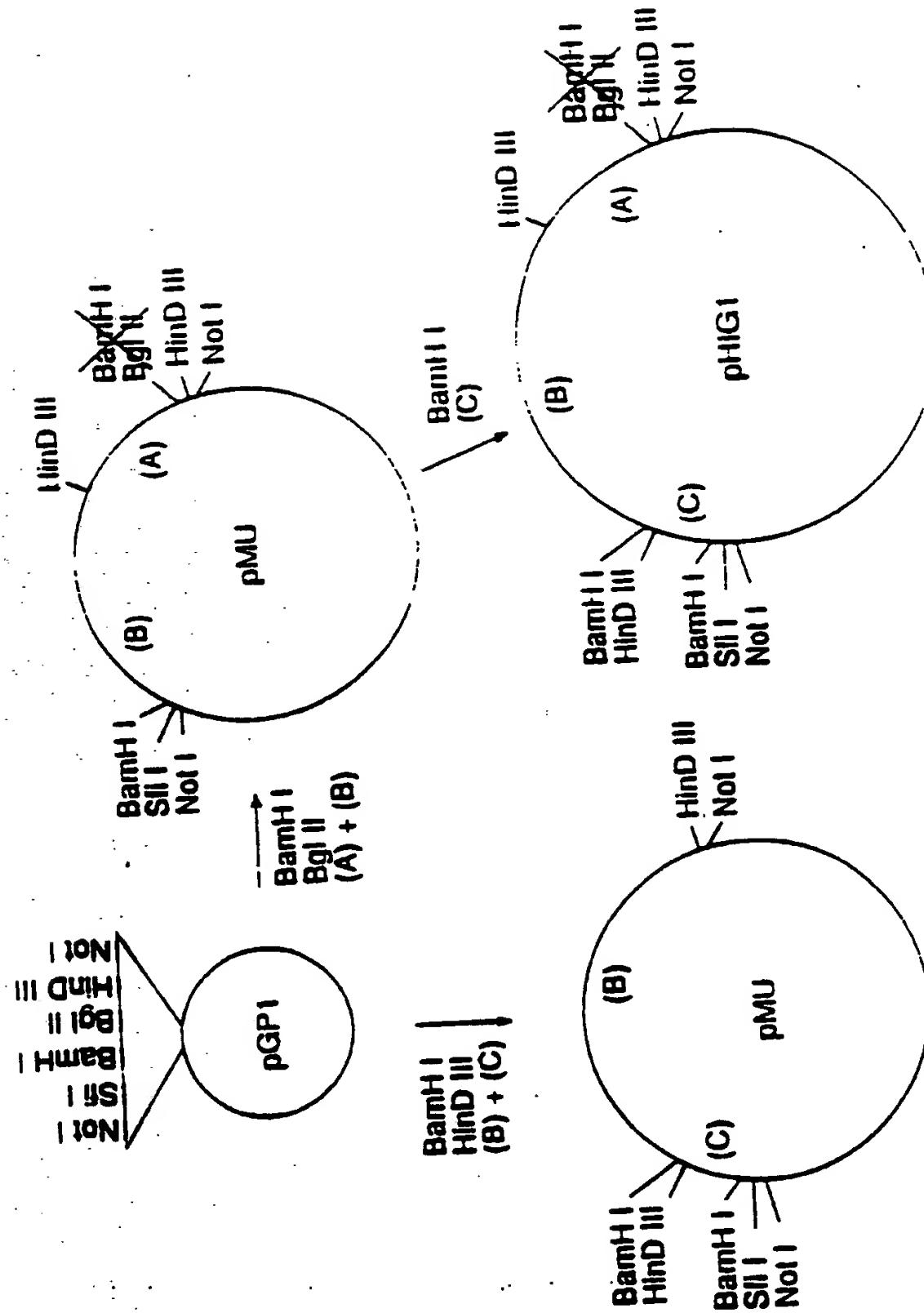
FIGURE 5

FIGURE 6

FIGURE 7

pGP1 POLYLINKER

FIGURE 8

HUMAN μ LOCUS

FIGURE 9

FIGURE 10

FIGURE 11

31.5 Kb Not I/Spe I INSERT

25 Kb Not I/Spe I INSERT

FIGURE 12

HUMAN D REGION
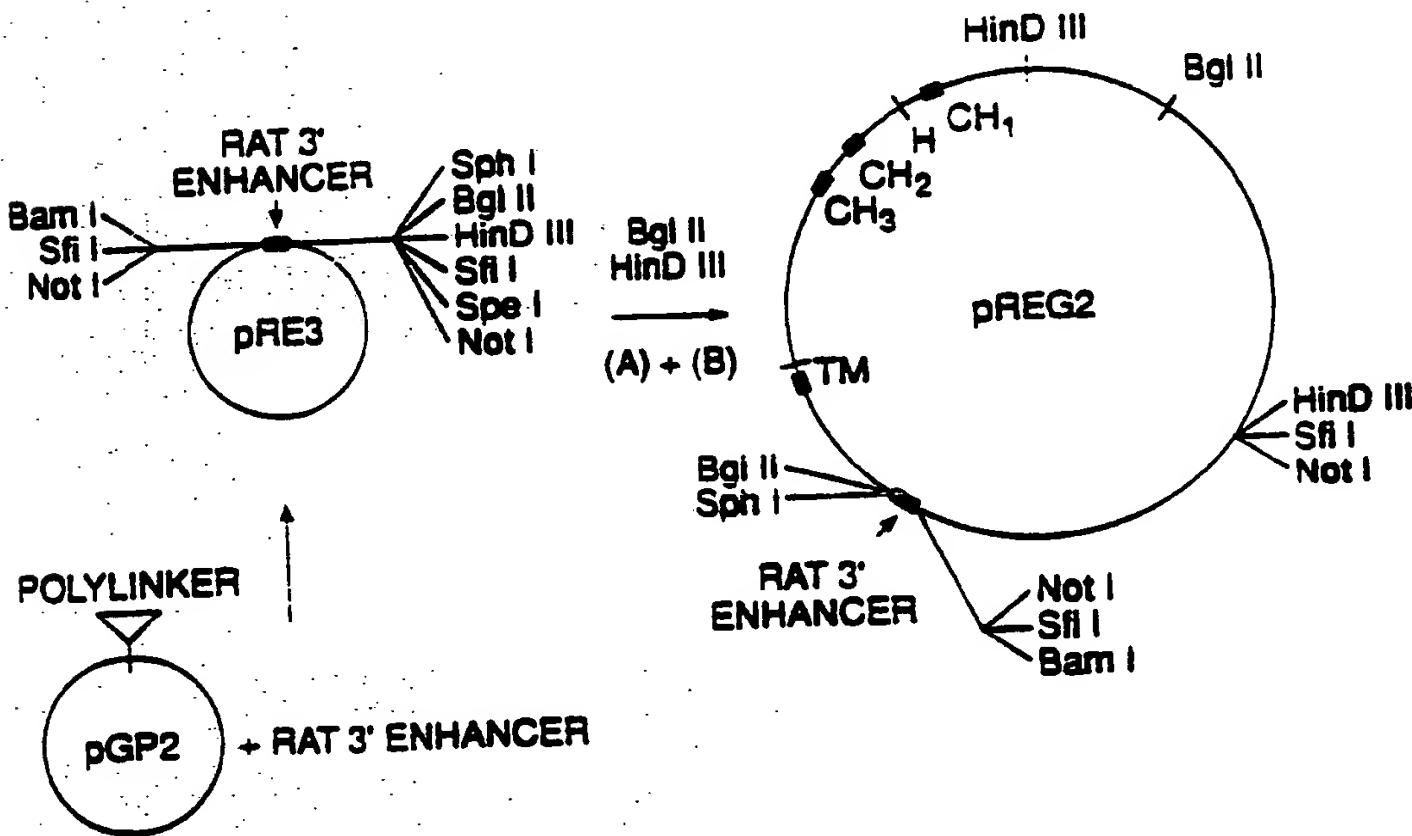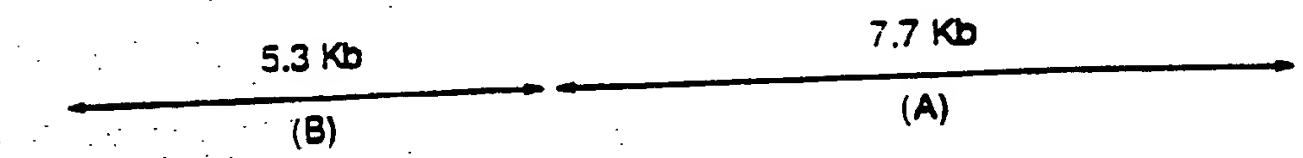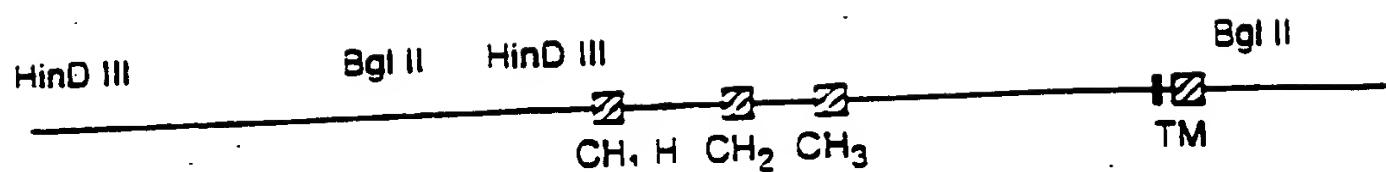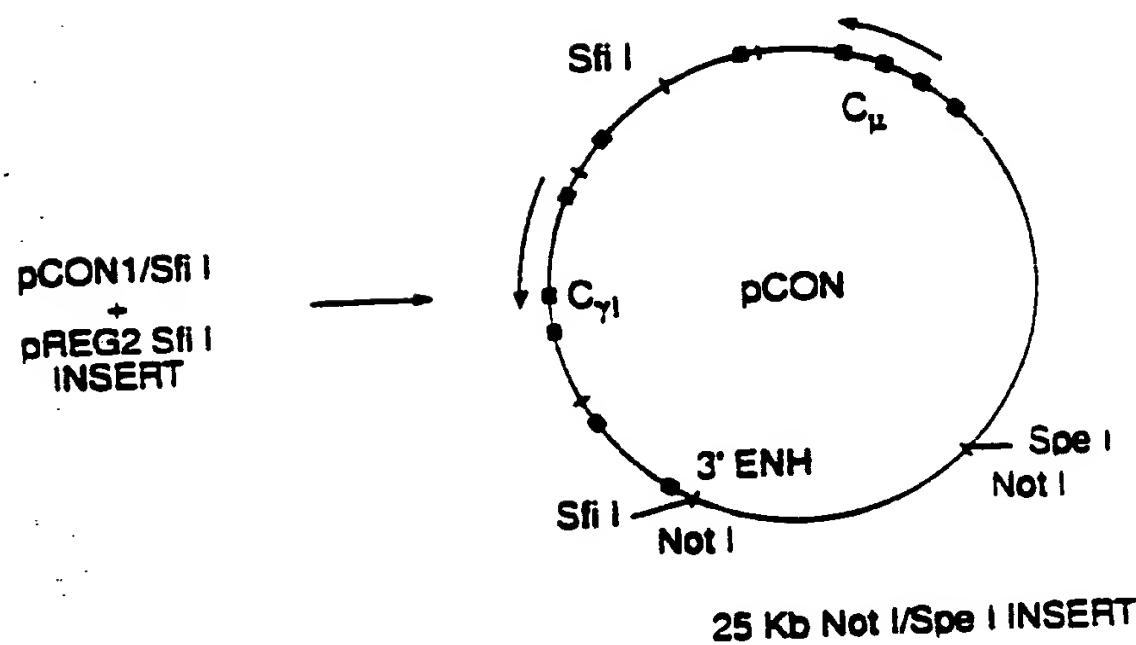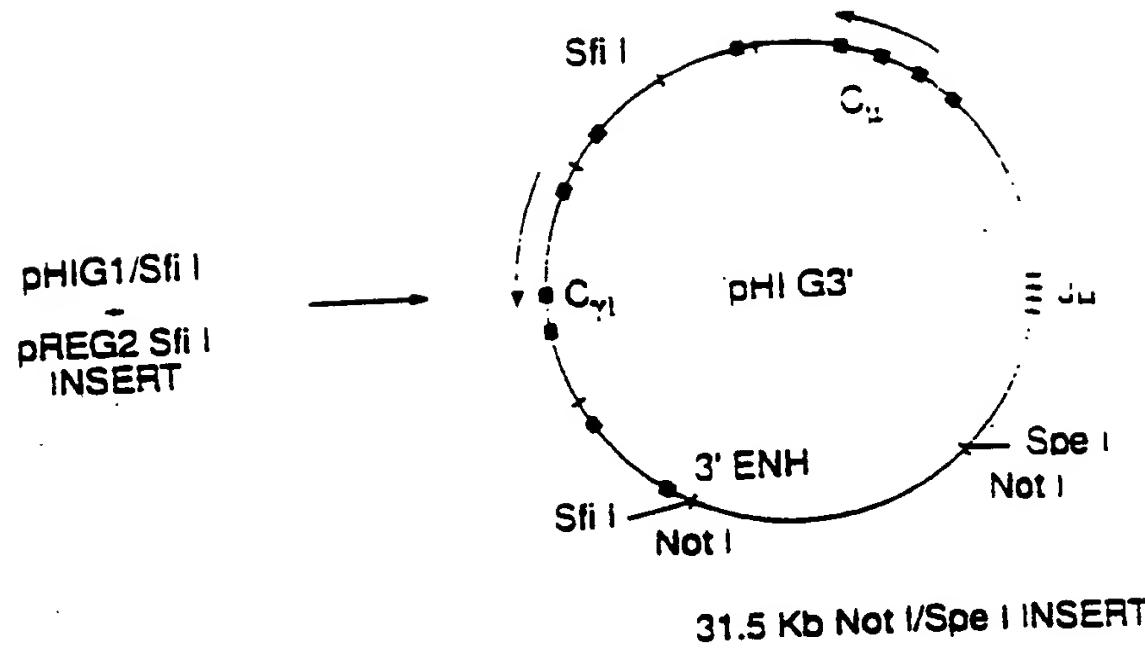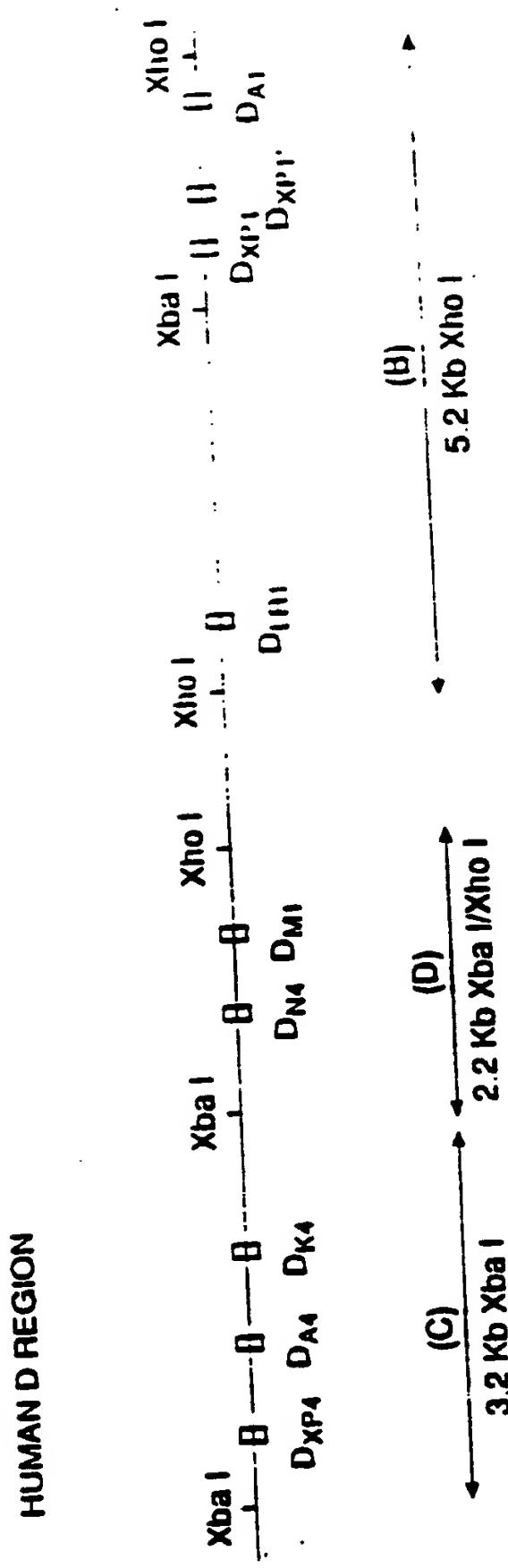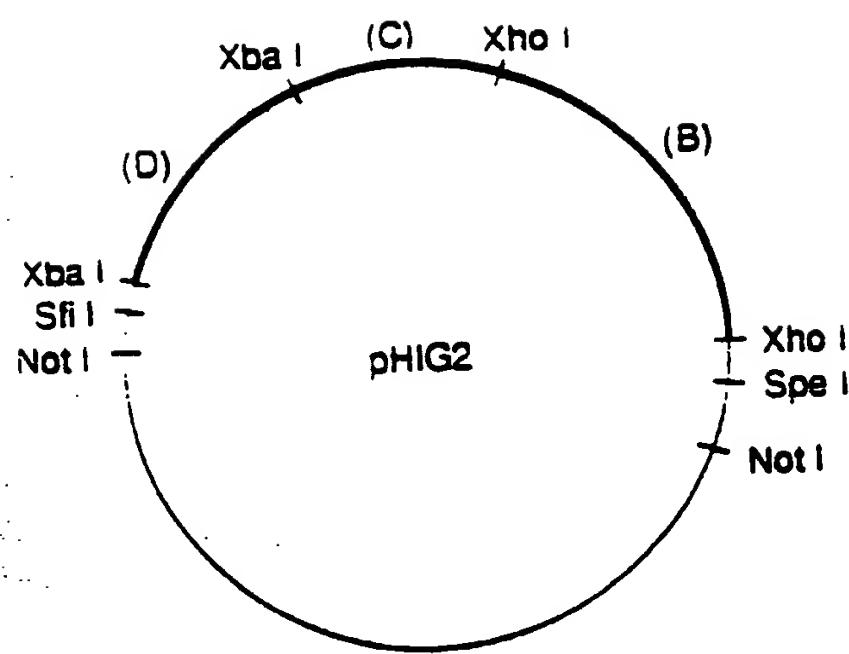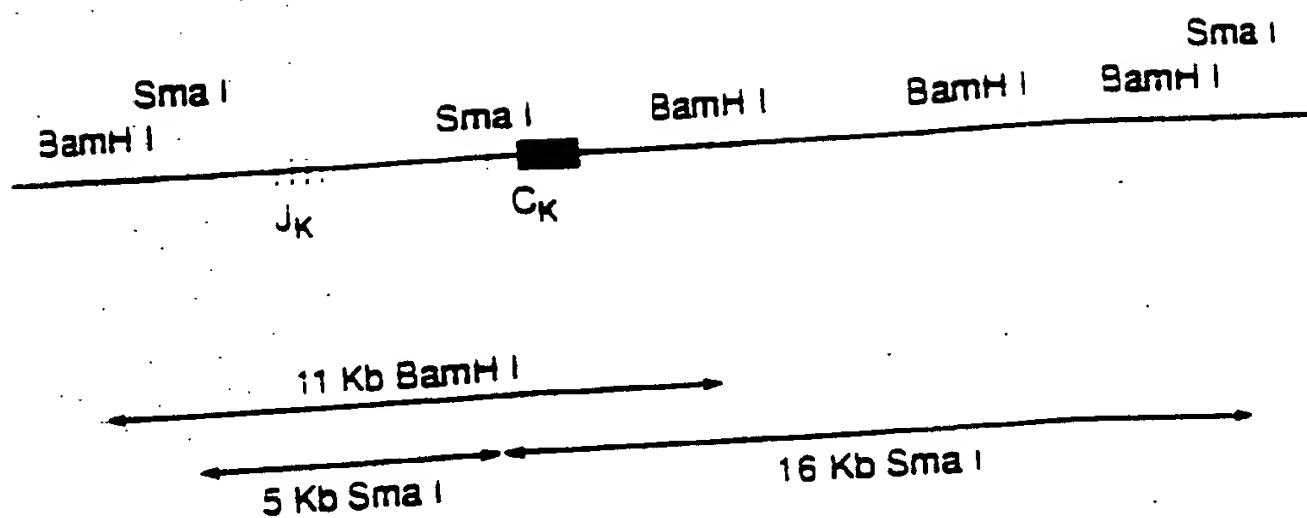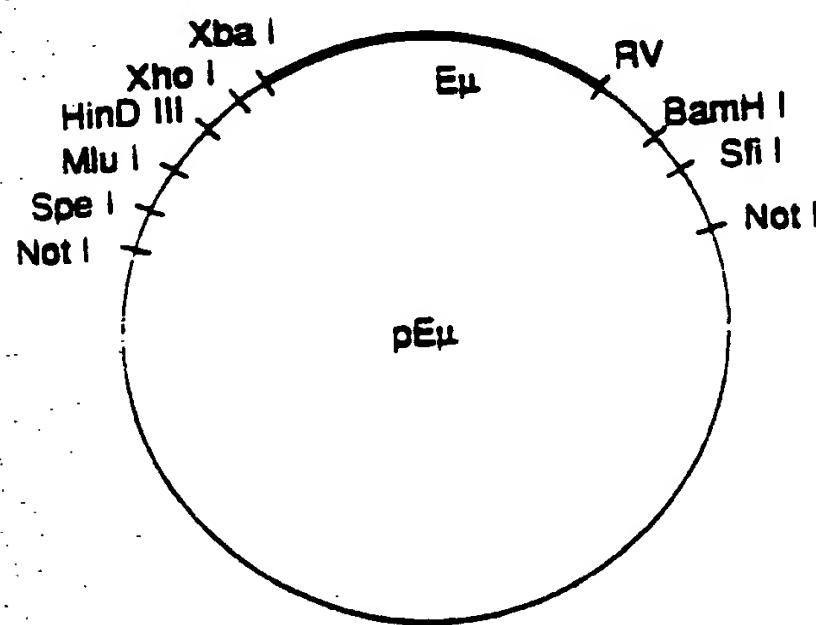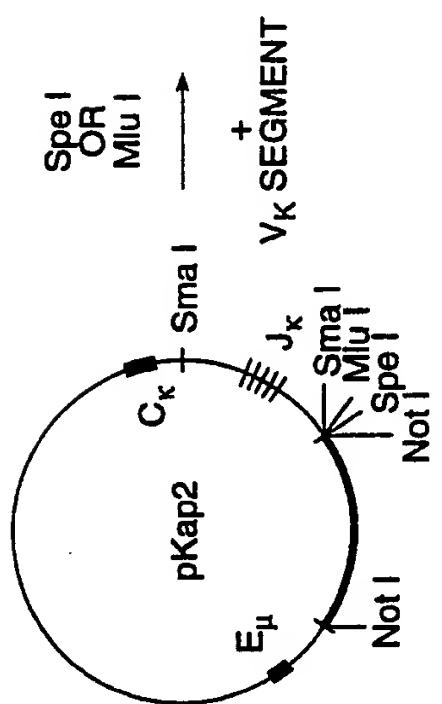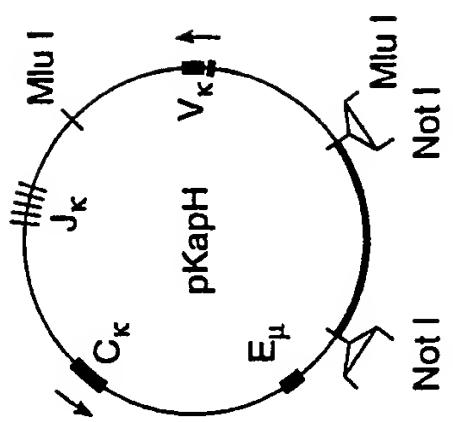
FIGURE 13

FIGURE 14
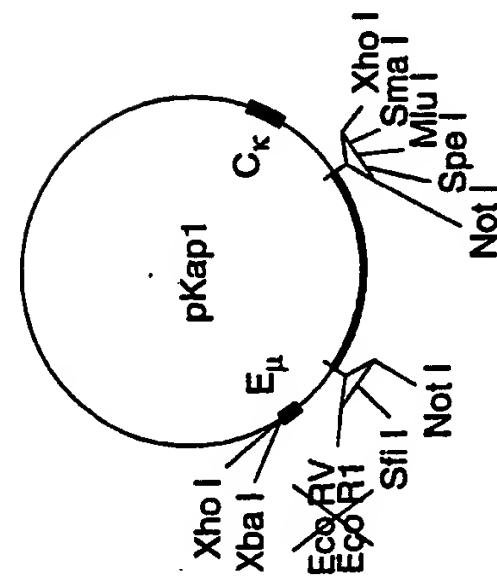
FIGURE 15

FIGURE 16

FIG. 17

MOUSE HEAVY CHAIN LOCUS
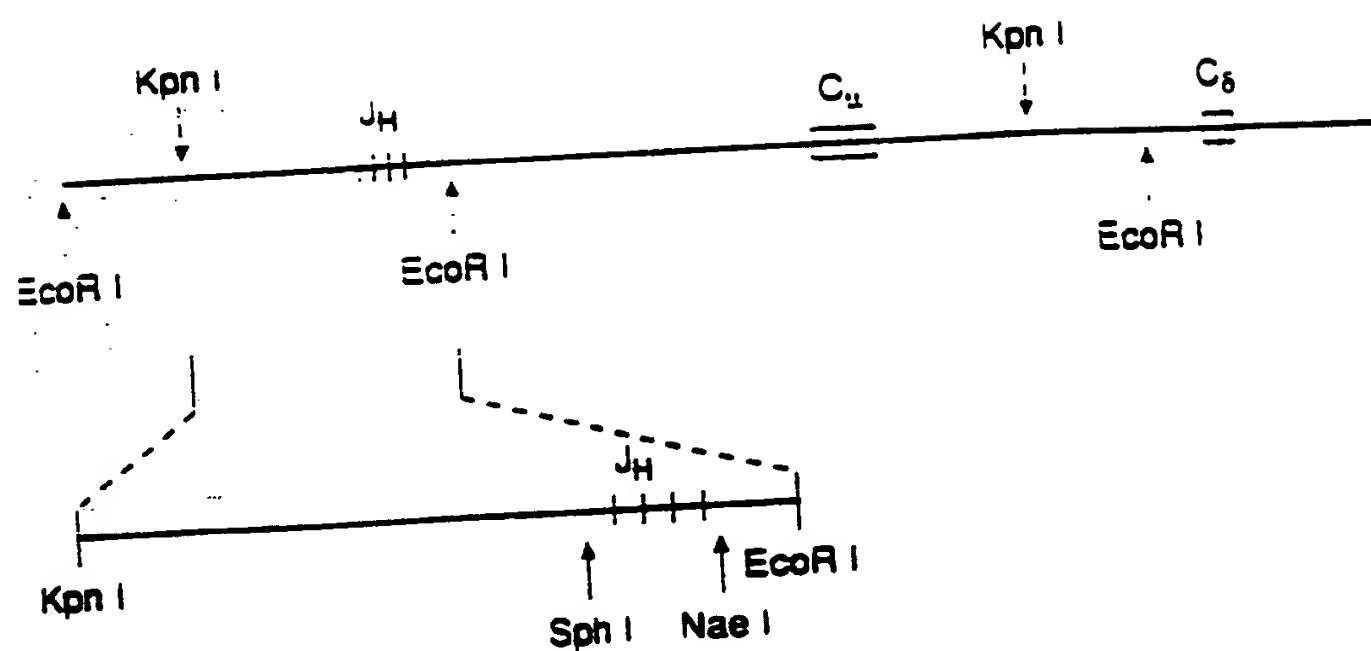
FIGURE 18a
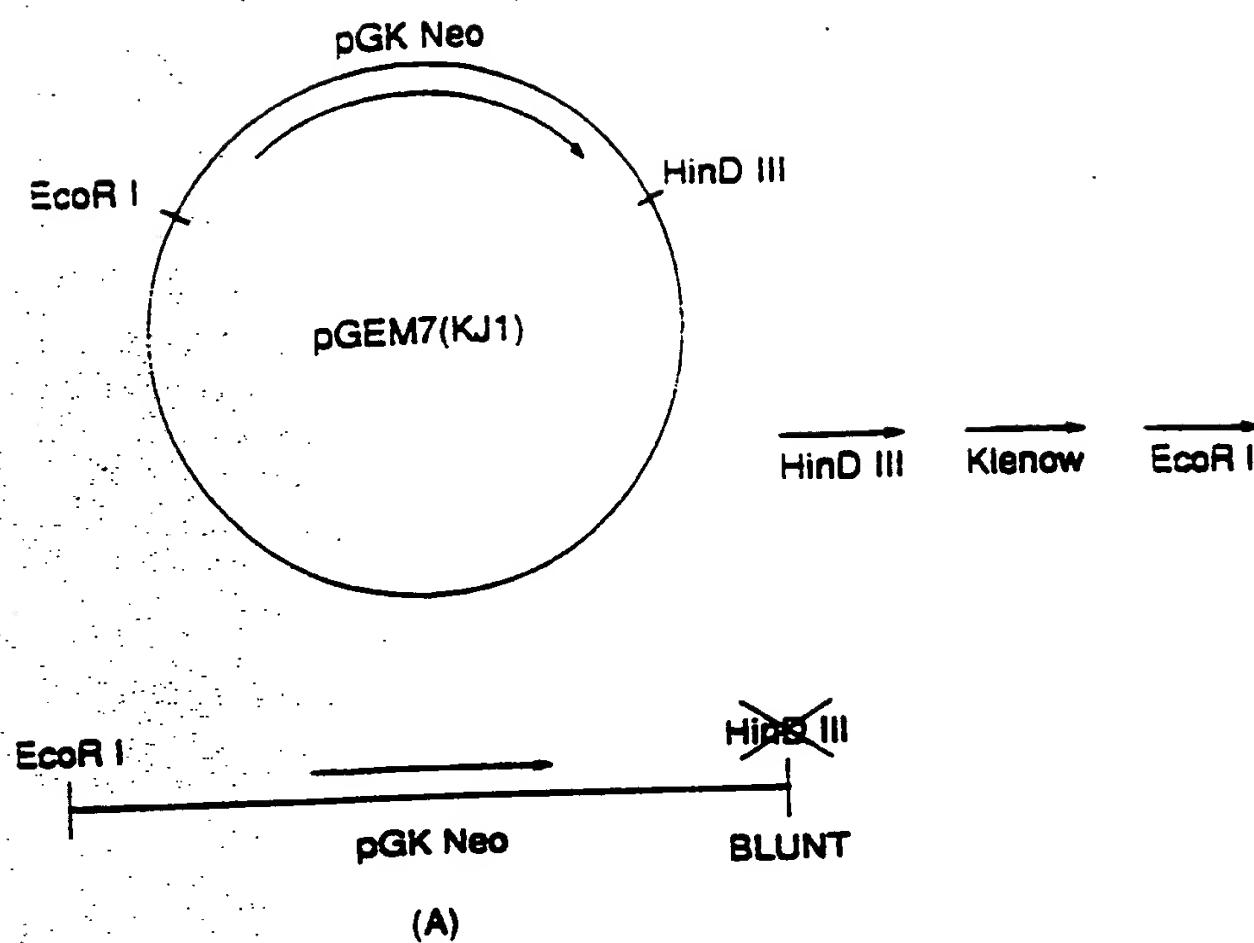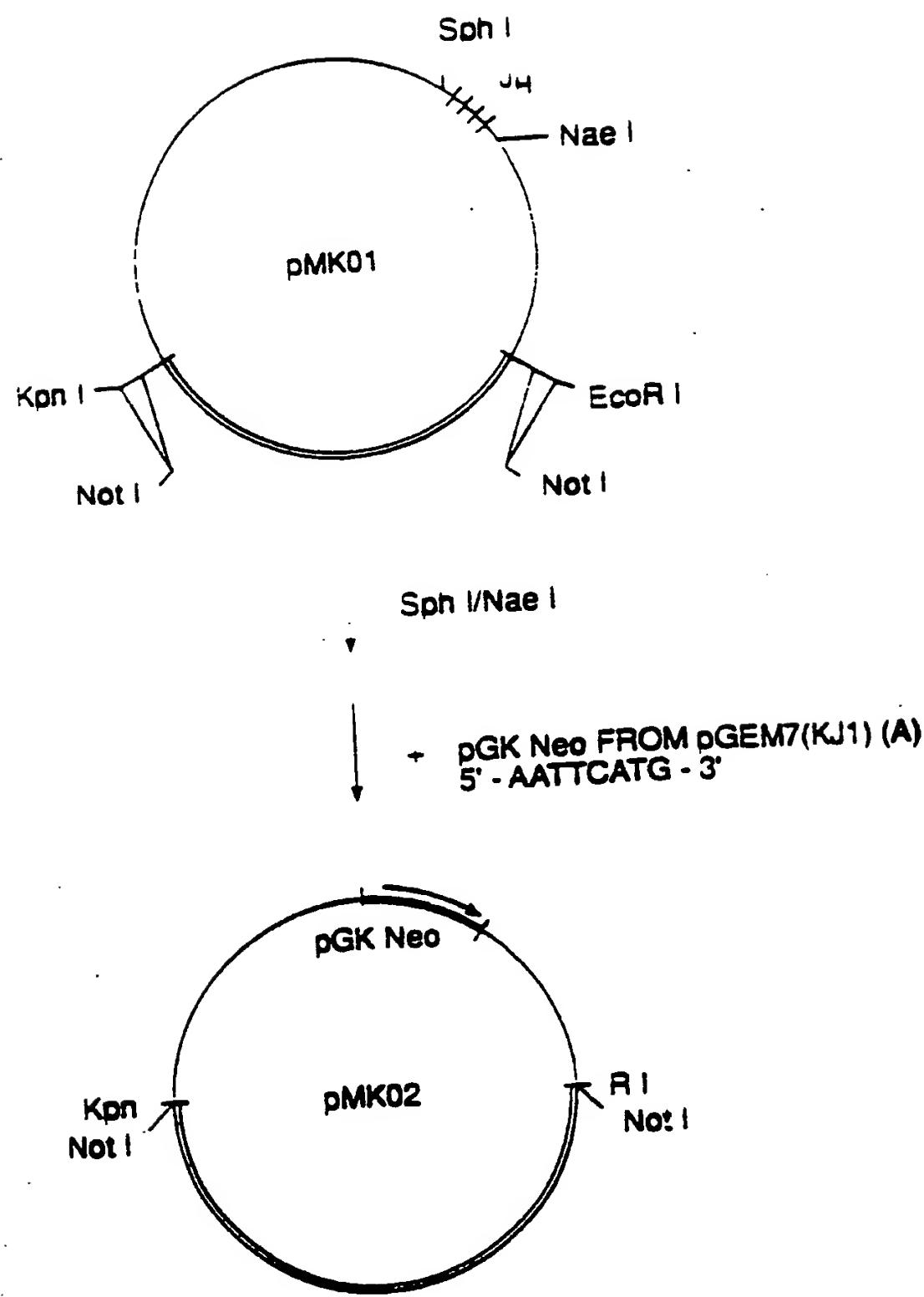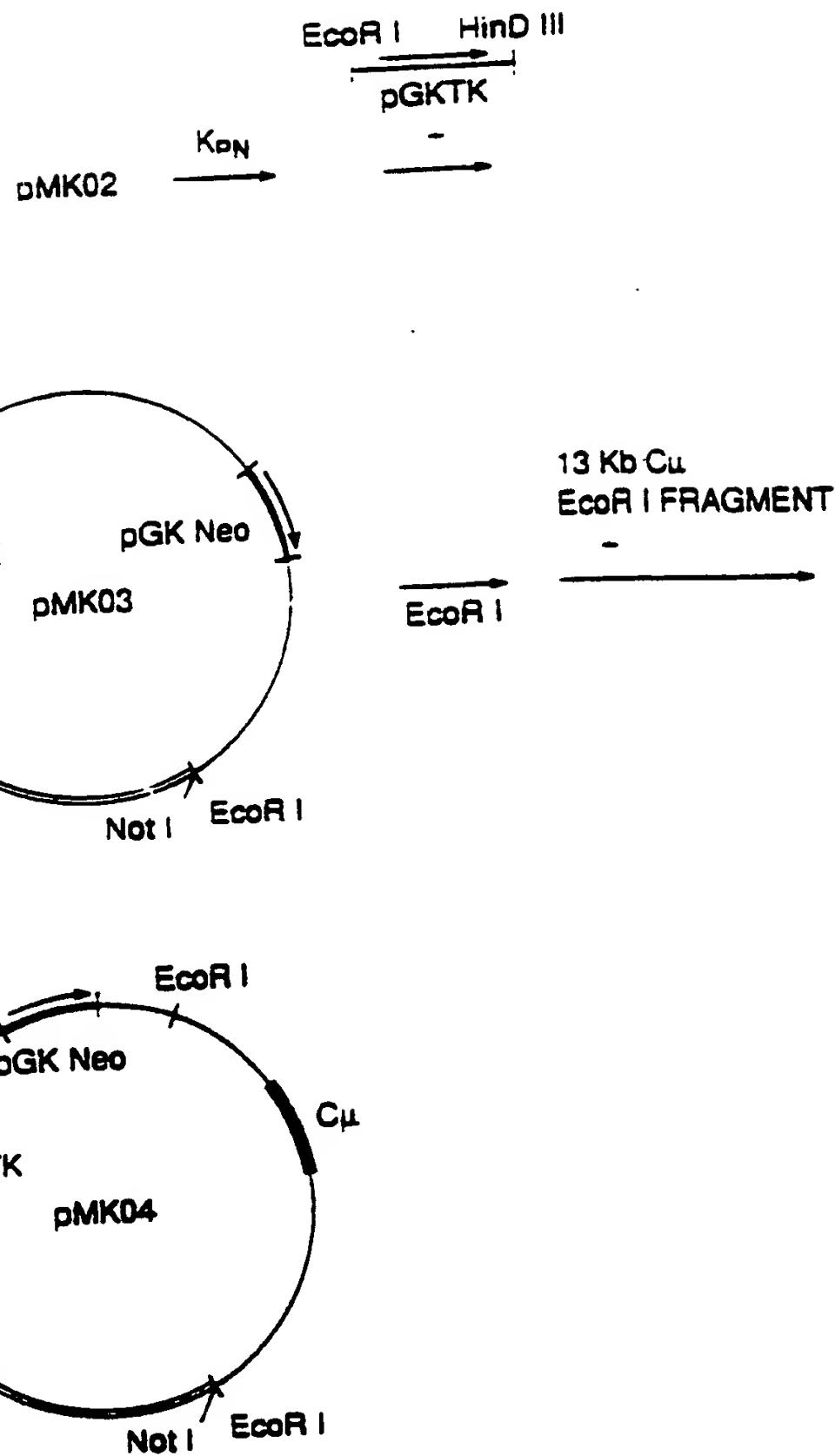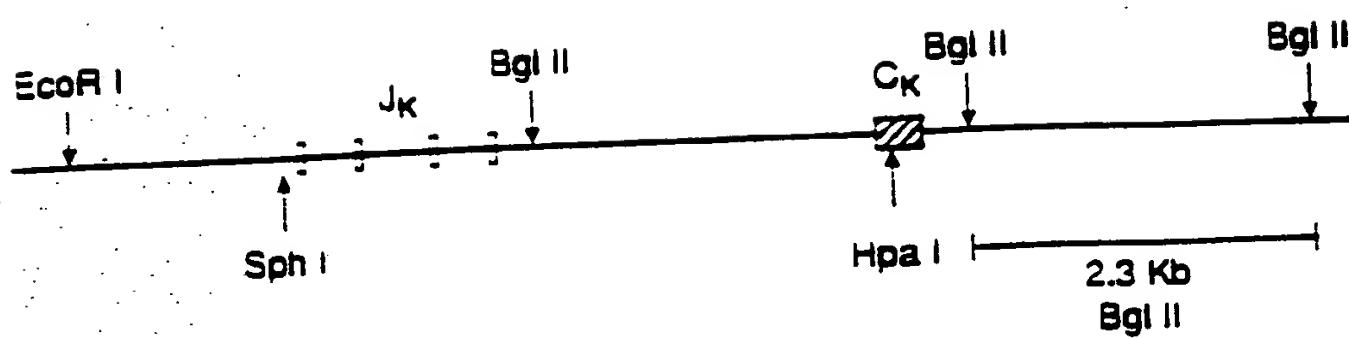
FIGURE 18b

FIGURE 18c

FIGURE 18d

MOUSE KAPPA GENE

FIGURE 19a
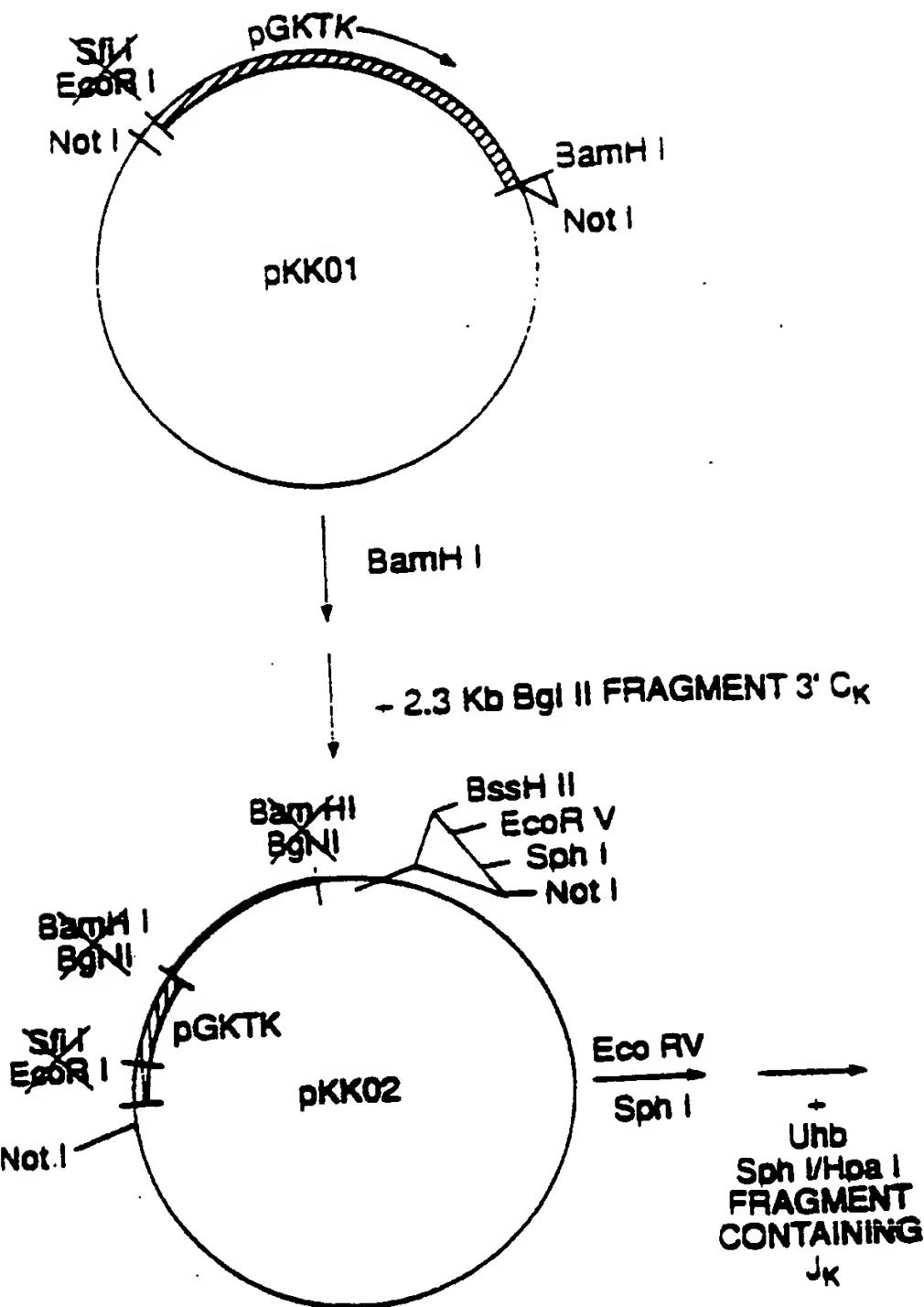
FIGURE 19b
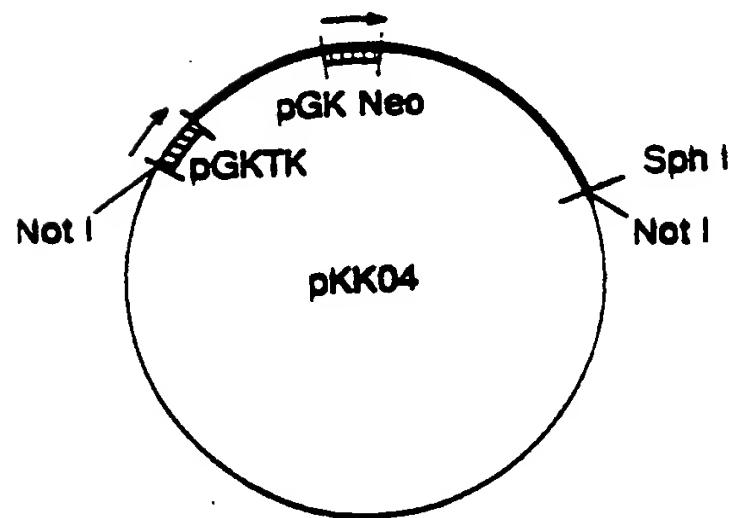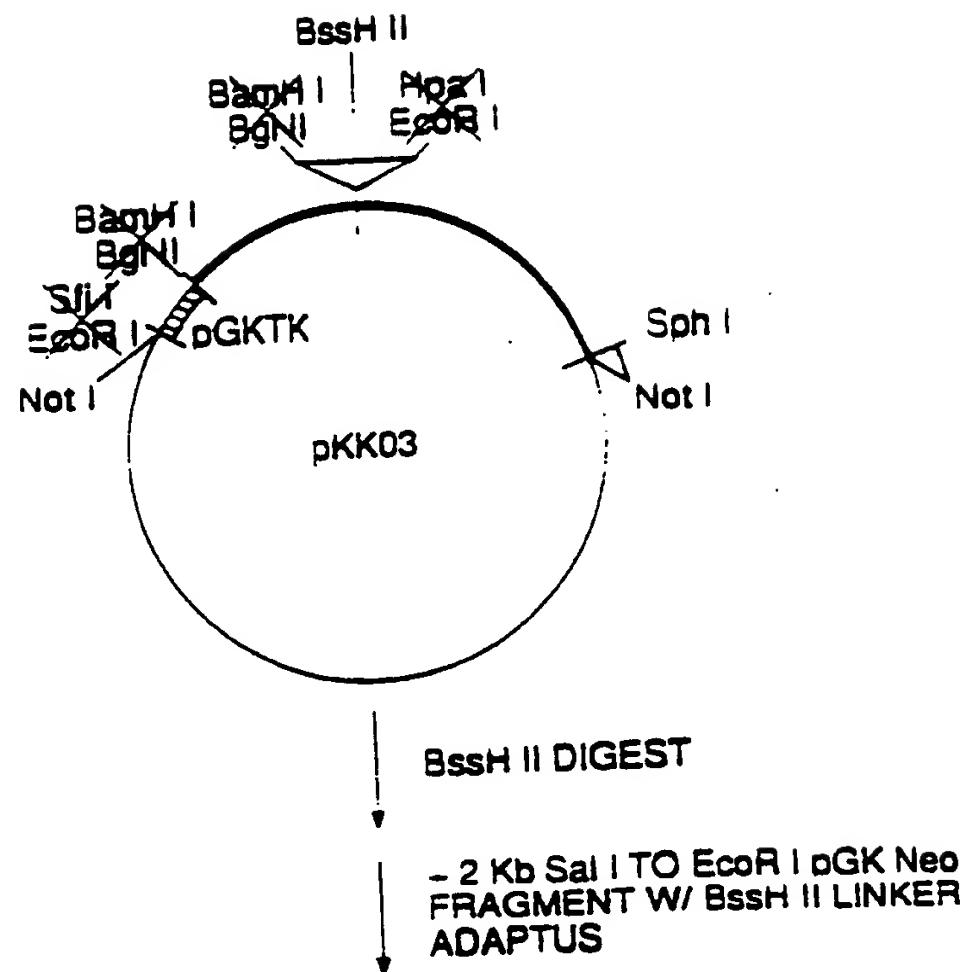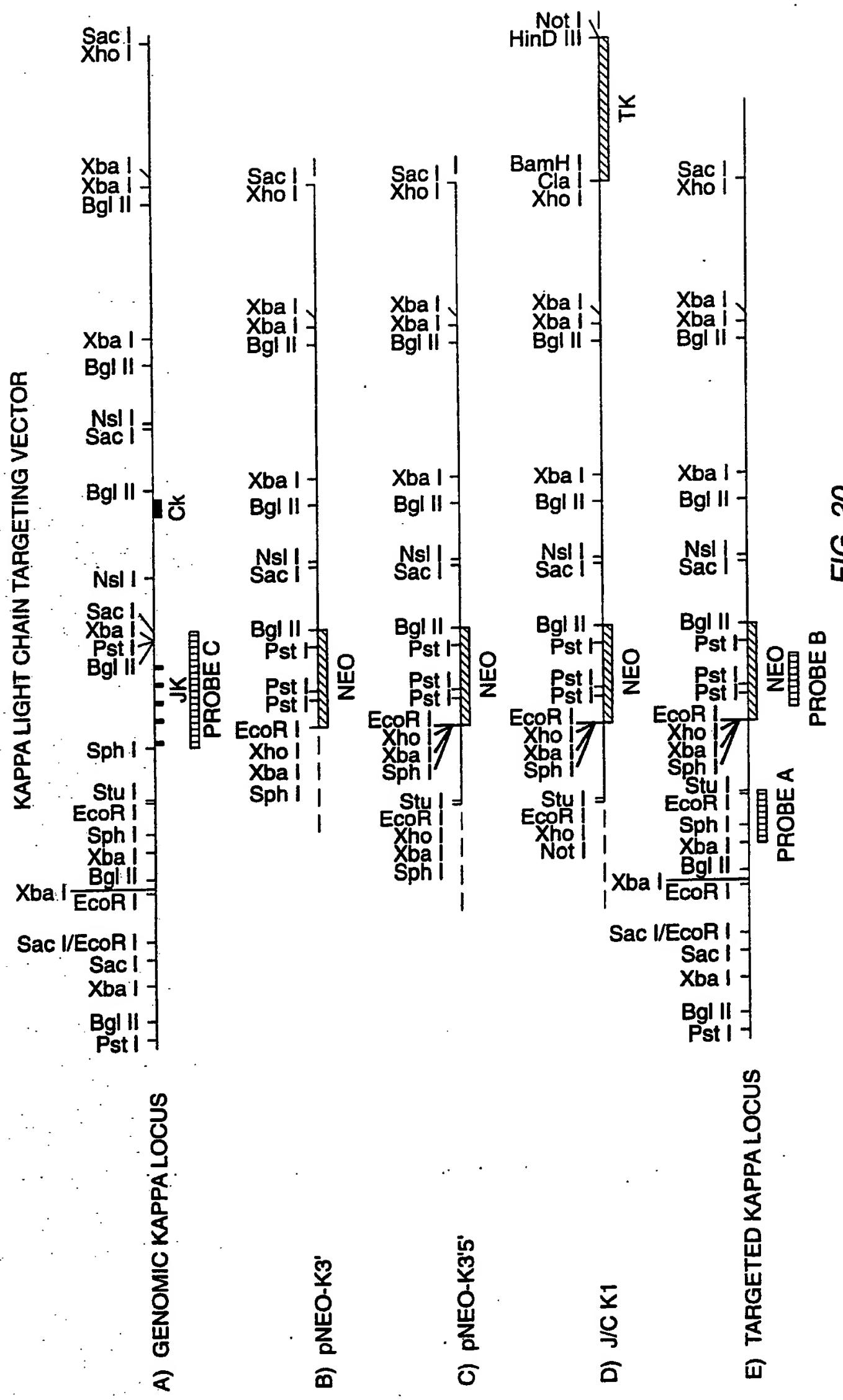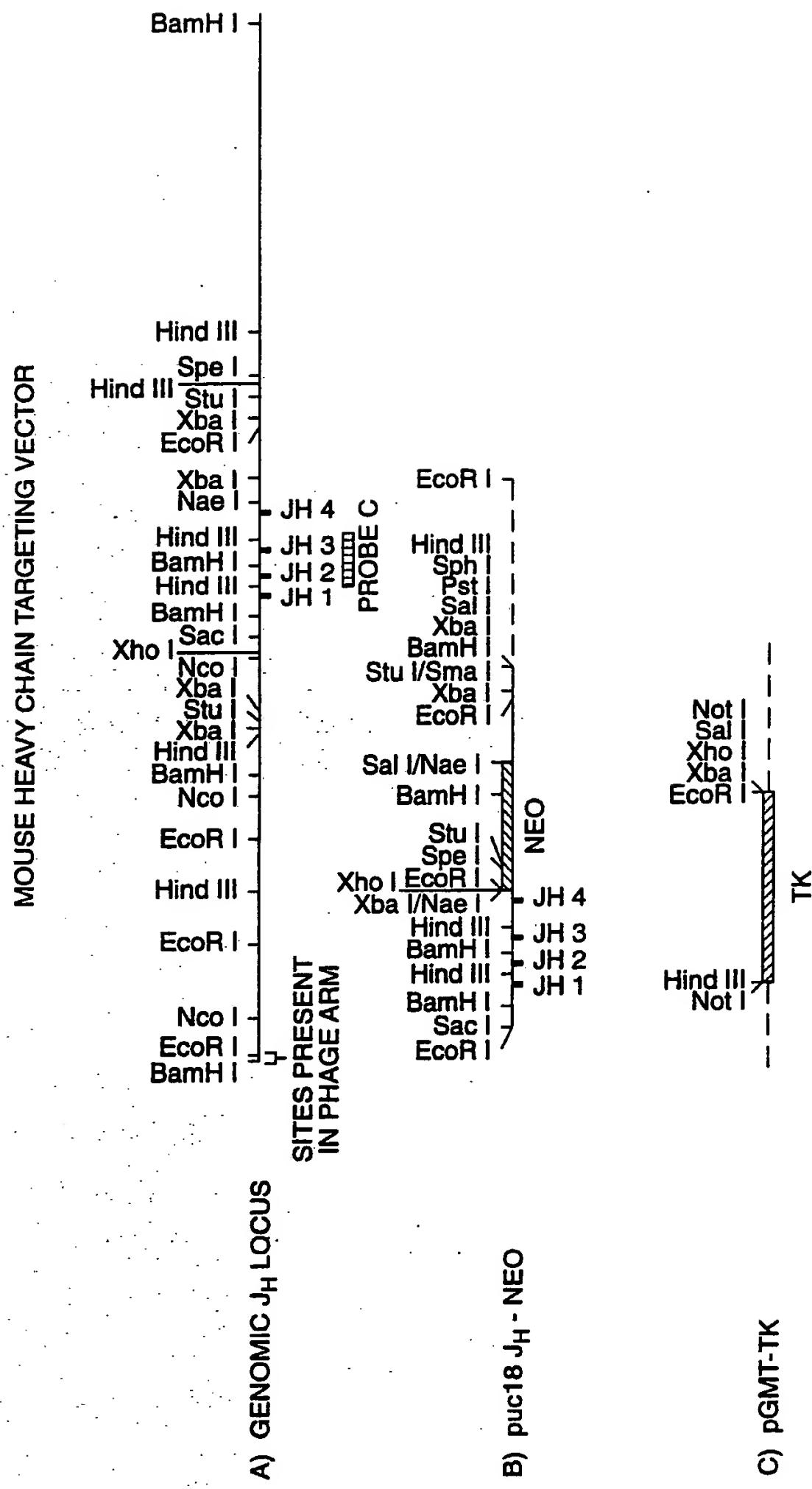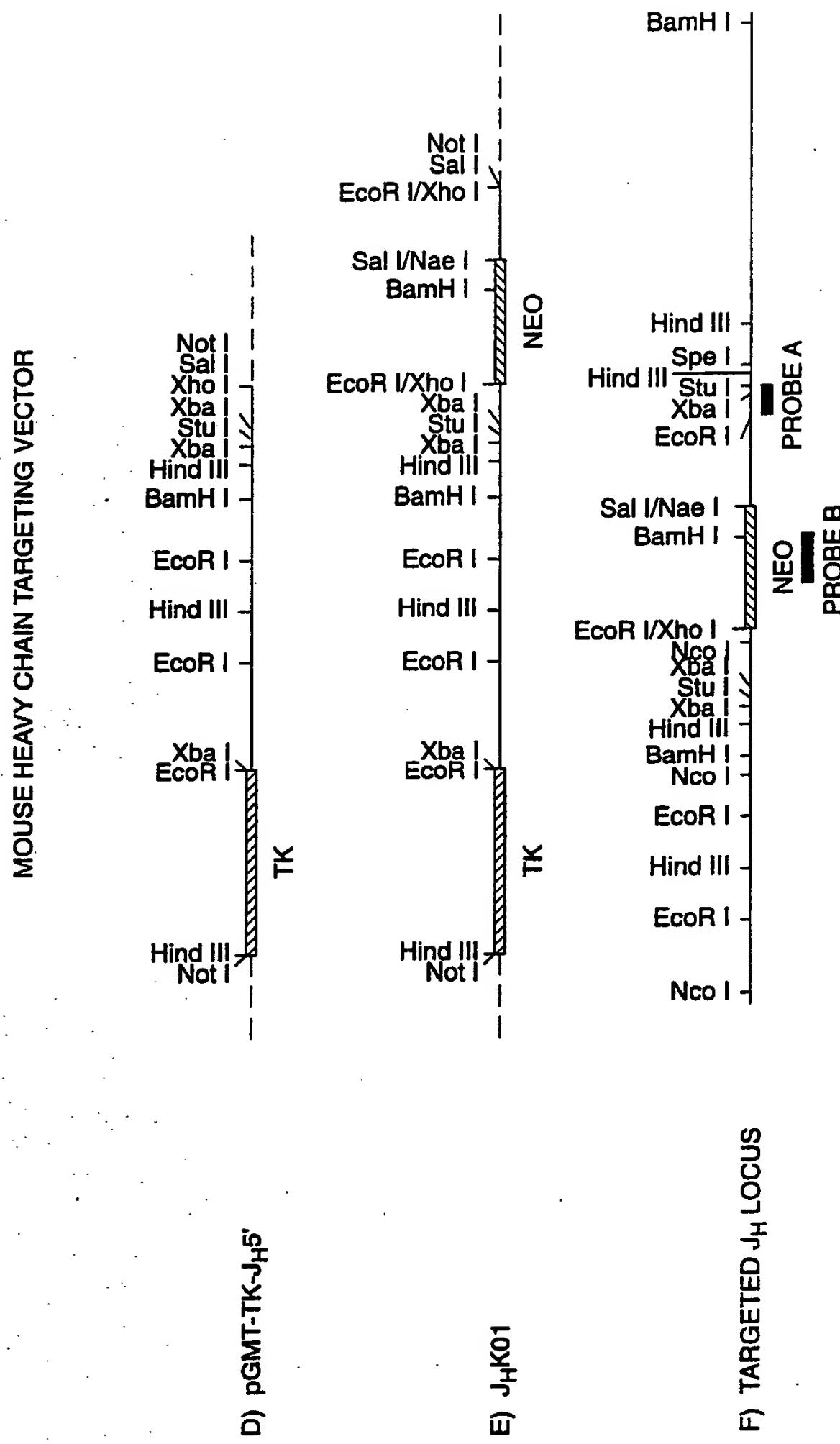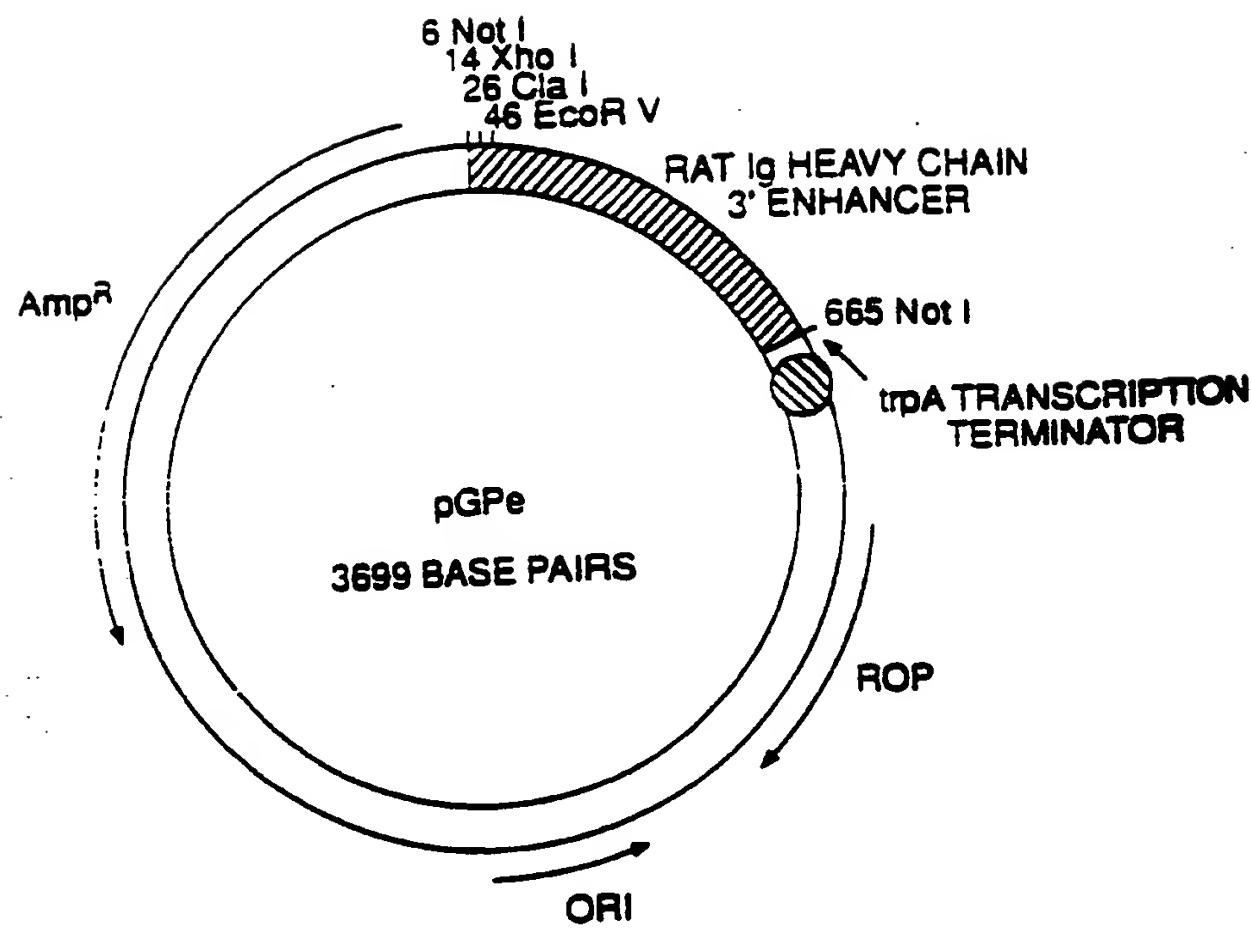
FIGURE 19c

FIG. 20

A) GENOMIC KAPPA LOCUS
B) pNEO-K3'
C) pNEO-K3'5'
D) J/C K1
E) TARGETED KAPPA LOCUS

FIG. 21a

A) GENOMIC J_H LOCUS

B) puc18 J_H - NEO

C) pGMT-TK

MOUSE HEAVY CHAIN TARGETING VECTOR

SITES PRESENT IN PHAGE ARM

FIG. 21b

D) pGMT-TK-J_H5'

E) J_HK01

F) TARGETED J_H LOCUS

FIGURE 22

FIGURE 23

FIGURE 24

FIG. 25

FIGURE 26

FIGURE 27

FIGURE 28

Not I Xho I   J_K   C_K   Sal I   Not I

pJCK1
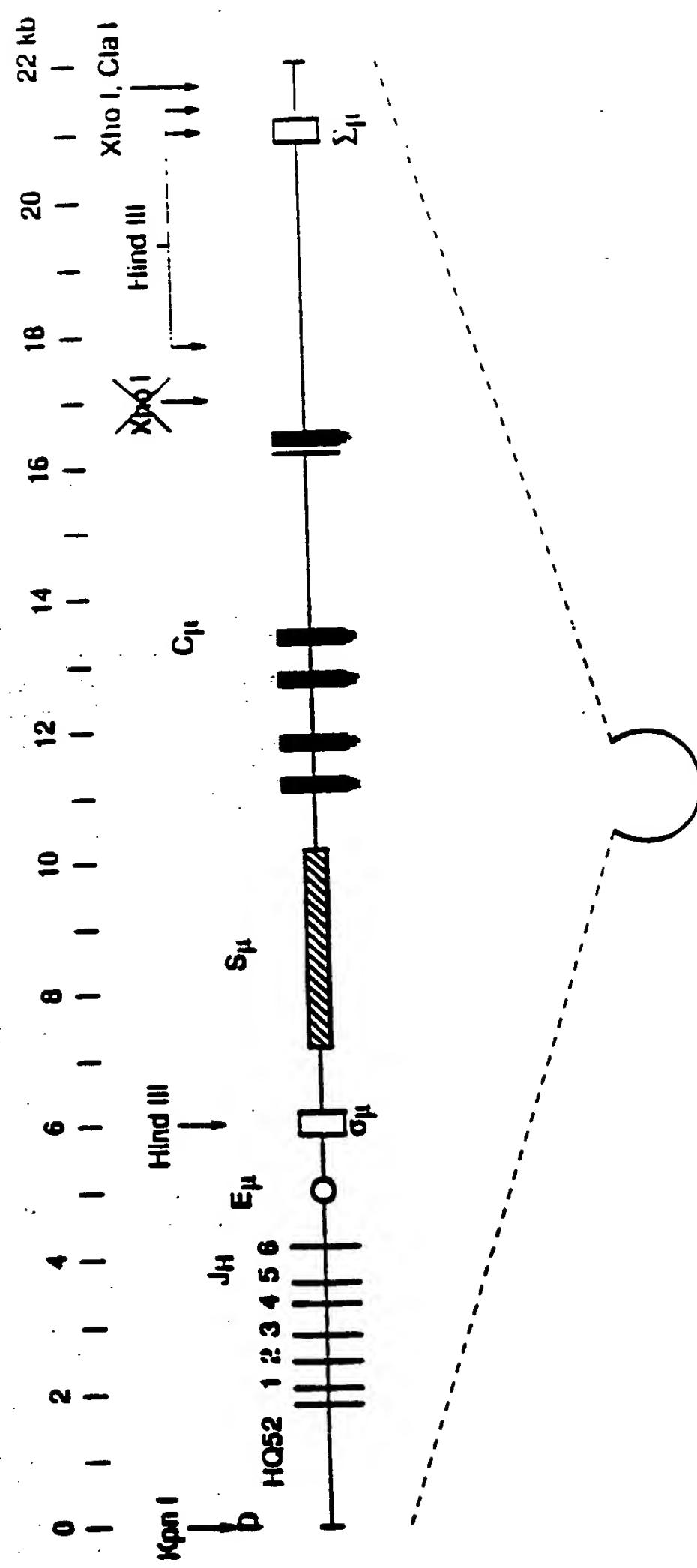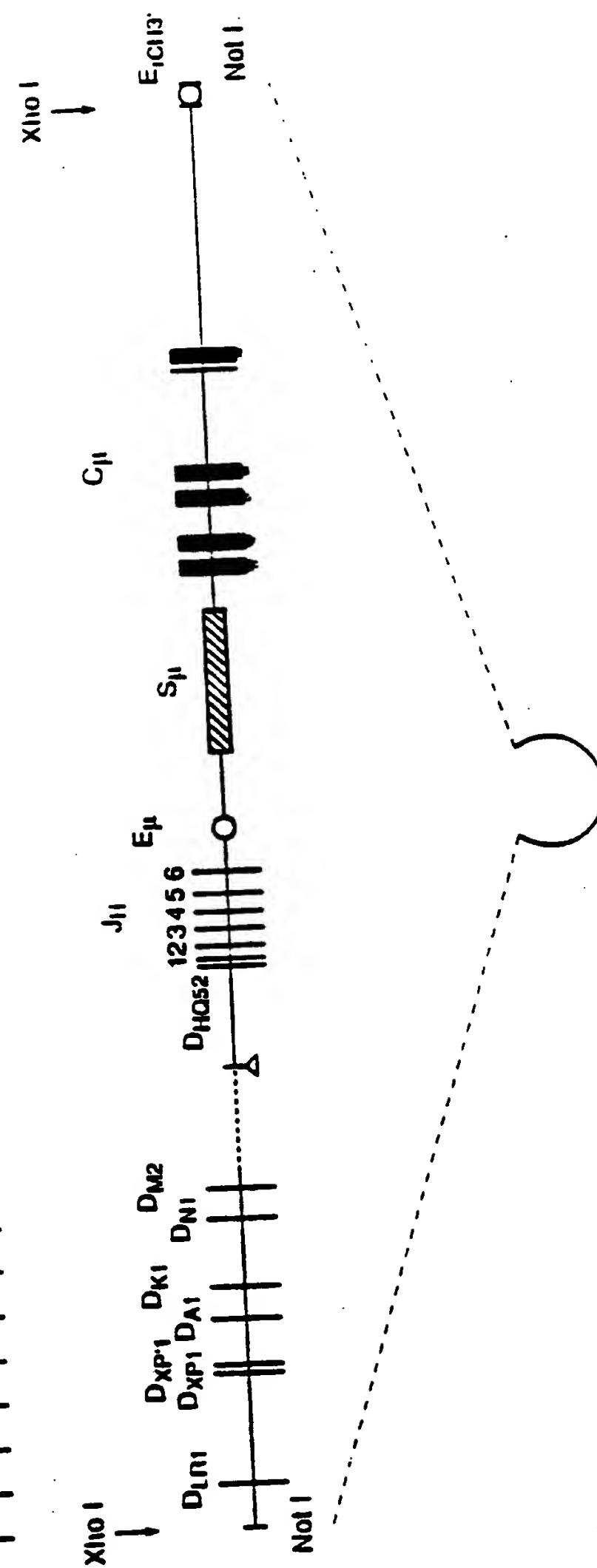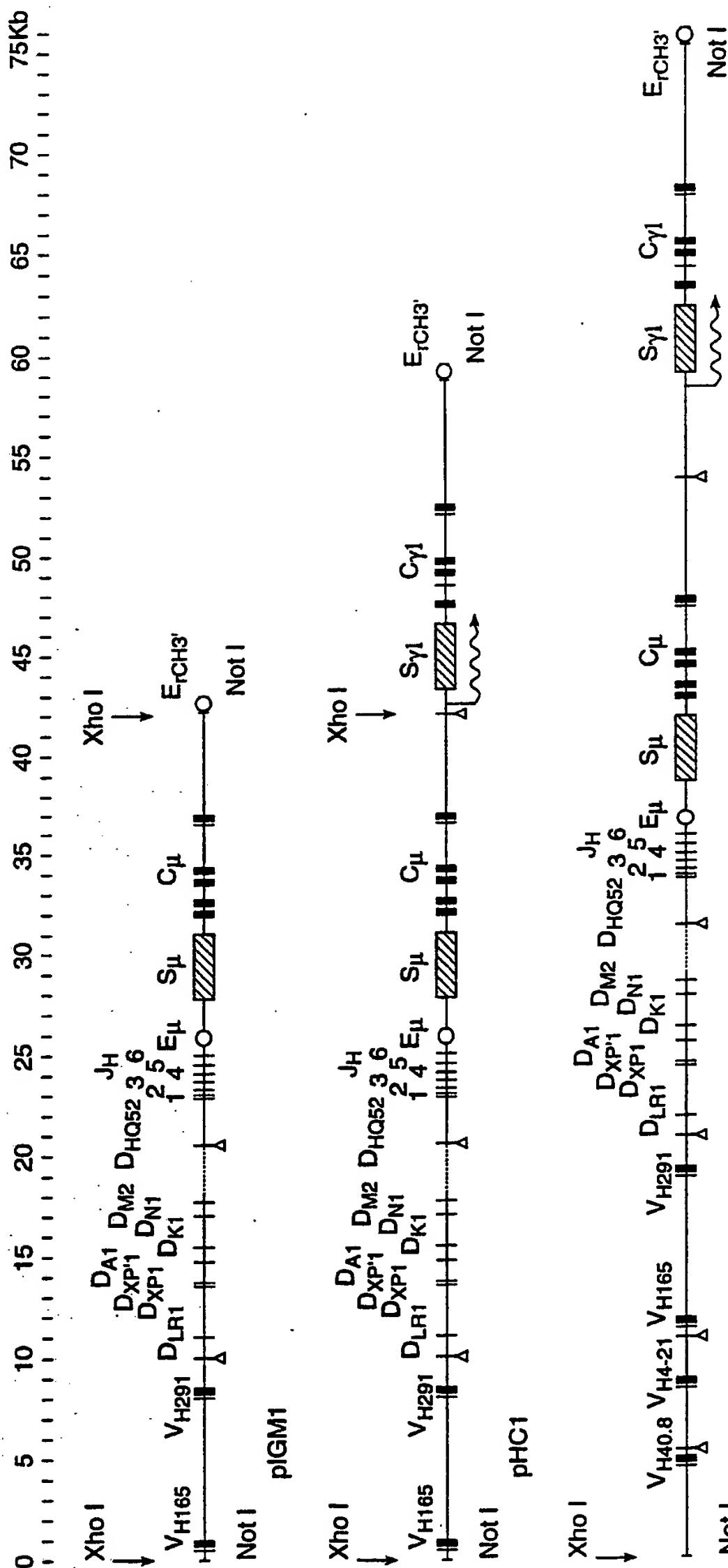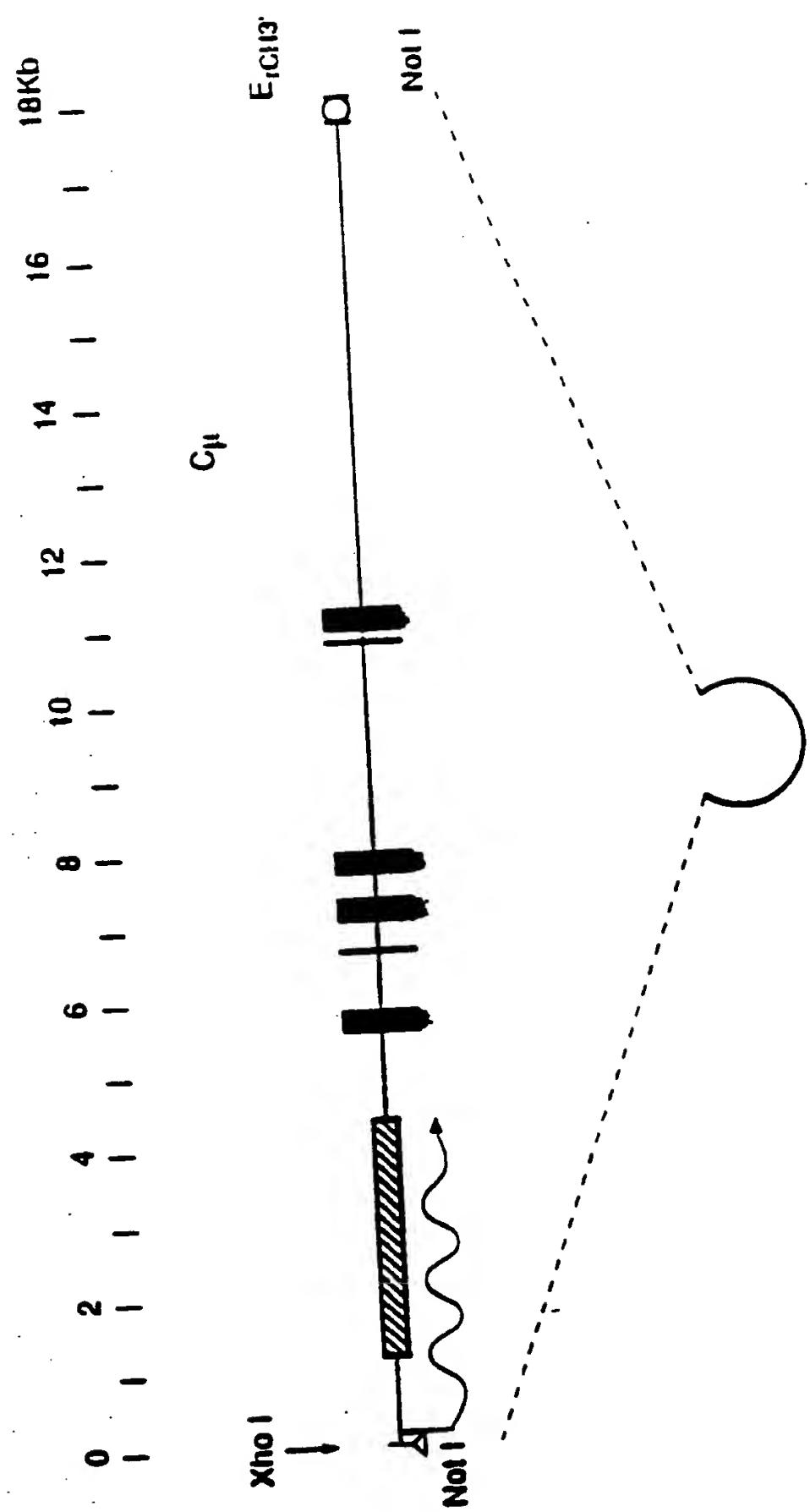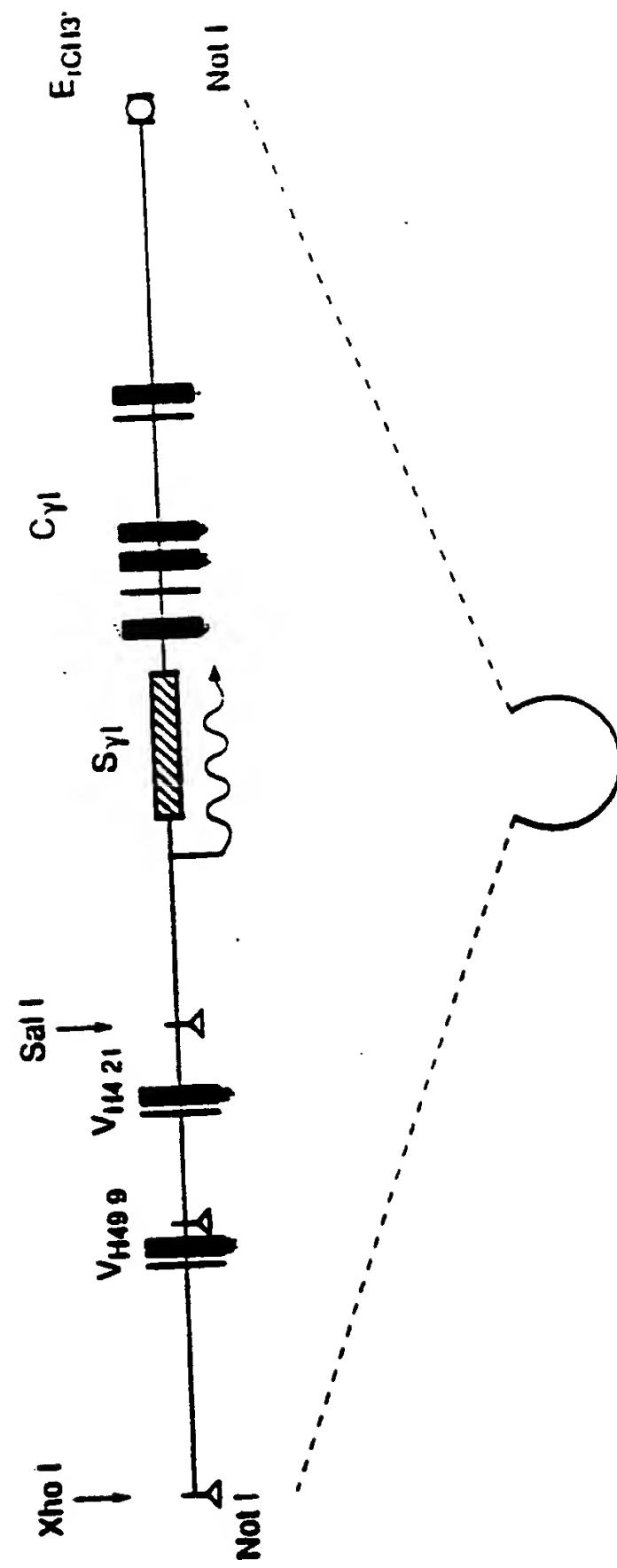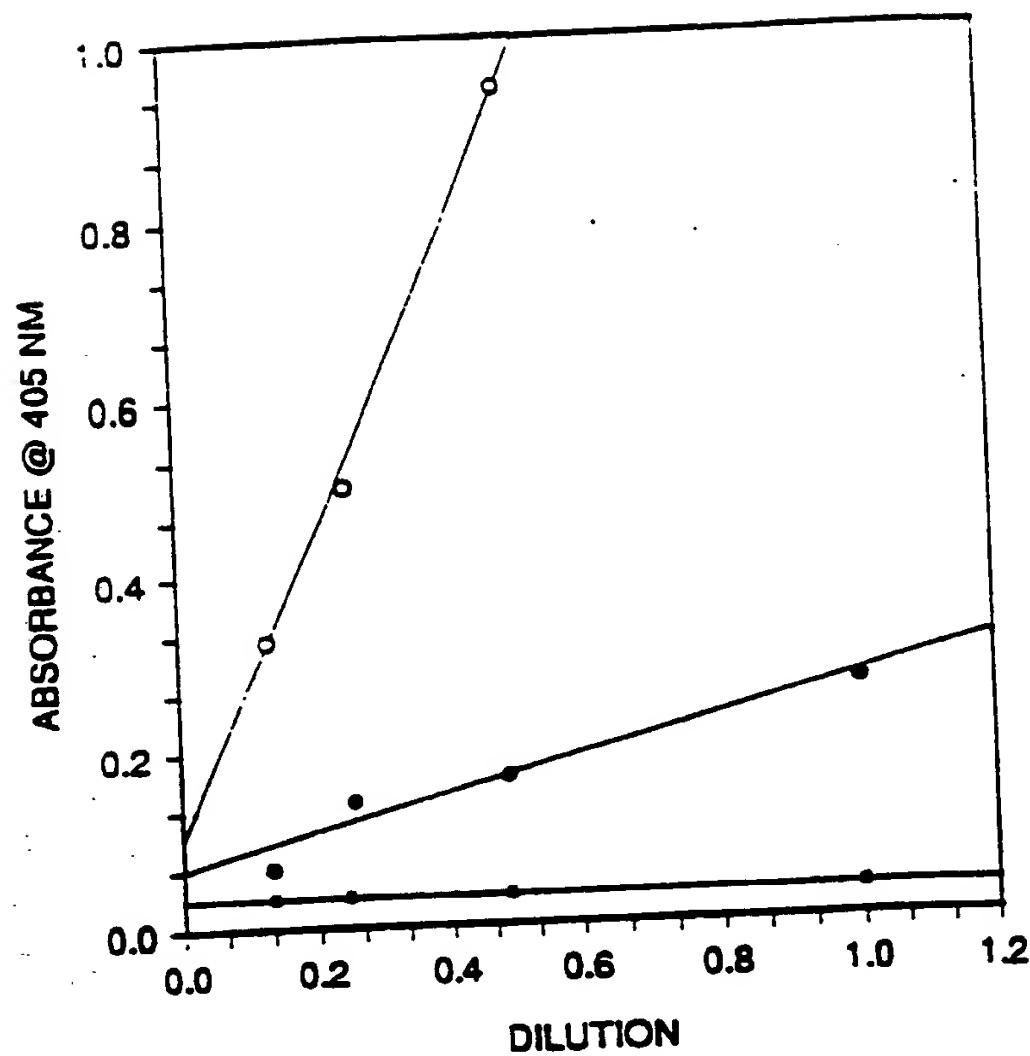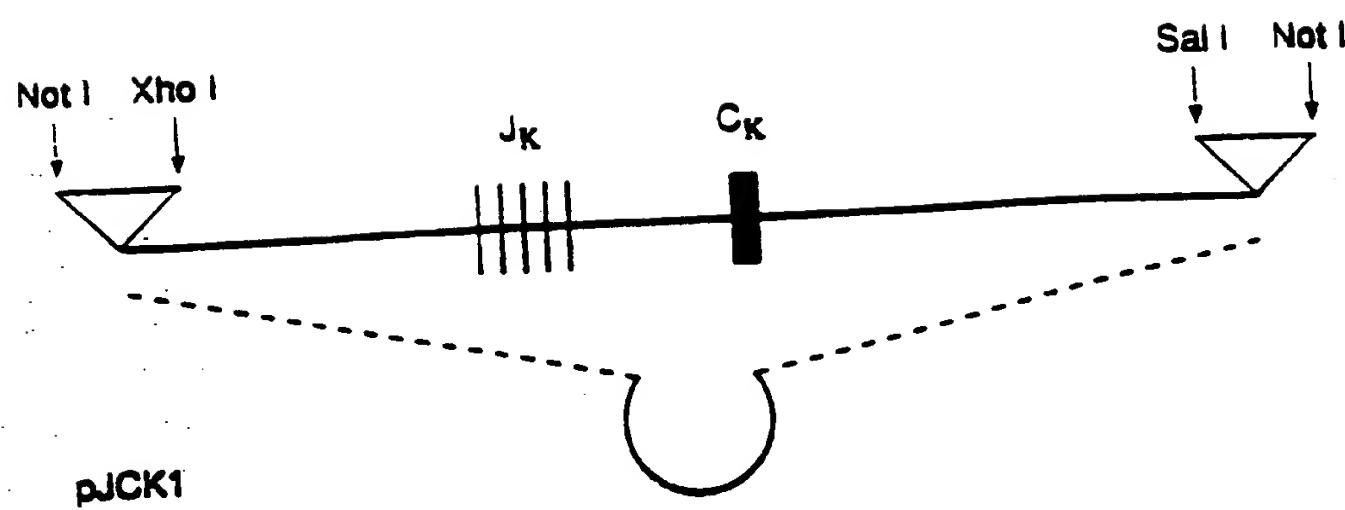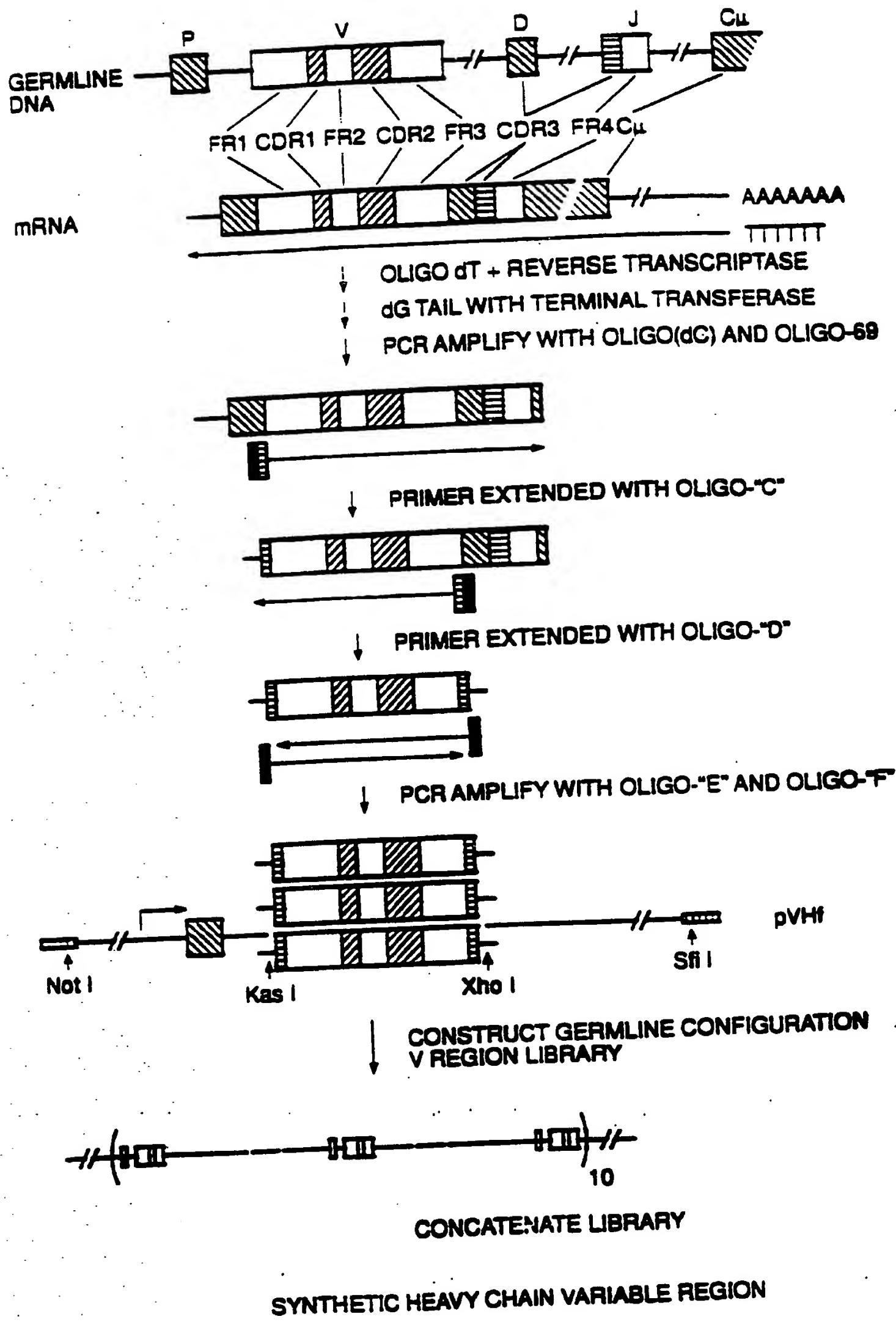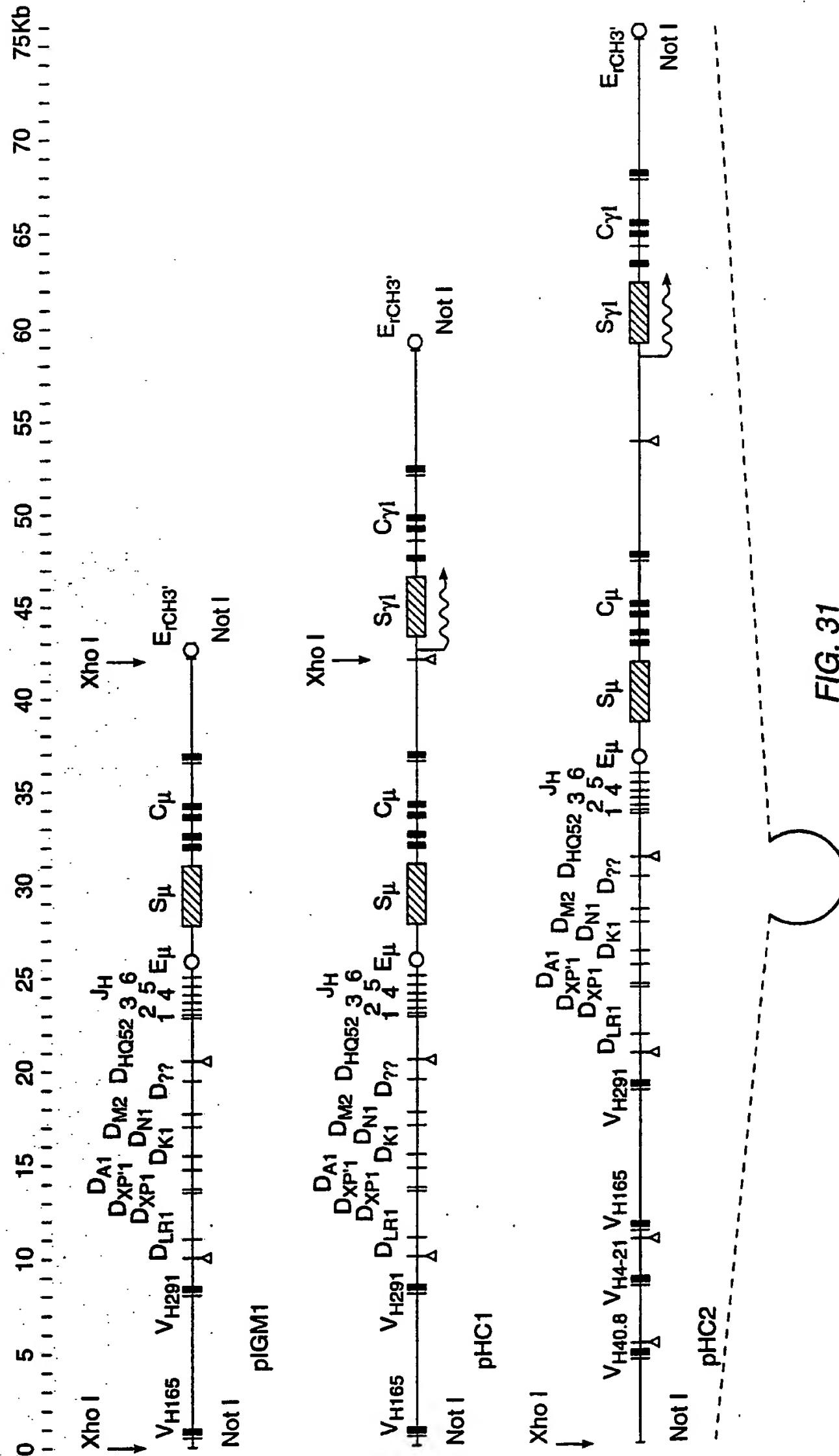
FIGURE 29

FIGURE 29

FIGURE 30

FIG. 31
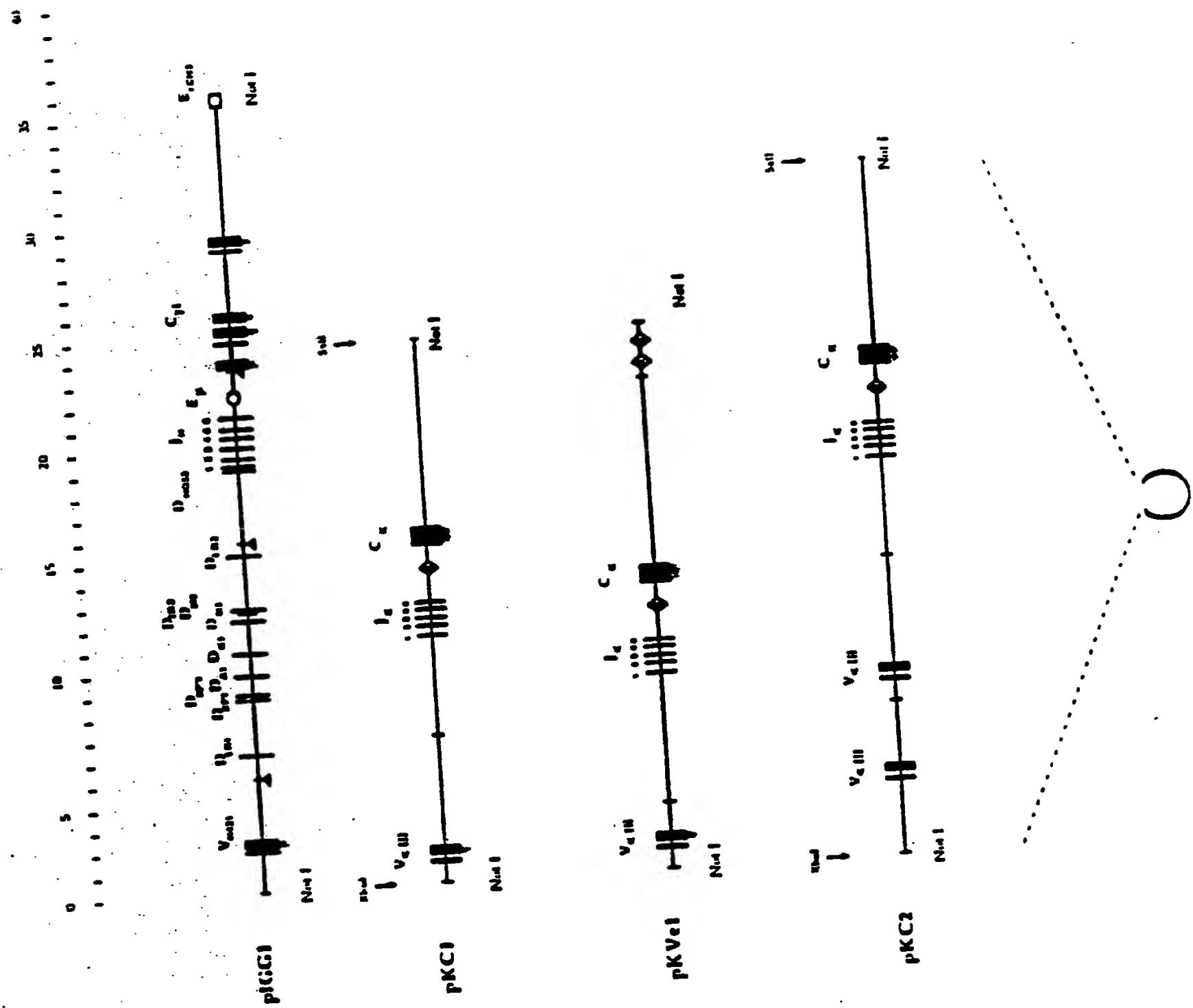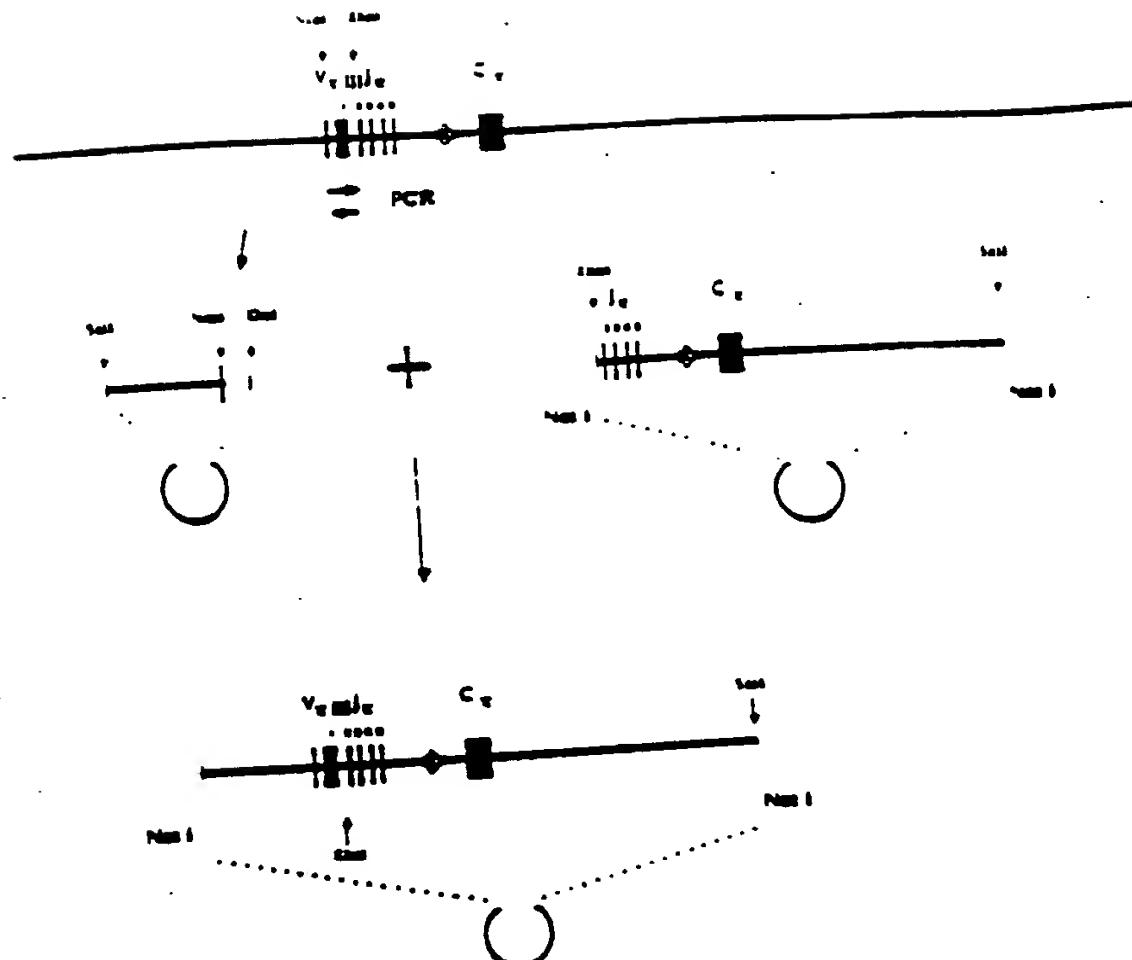
FIGURE 32

Expression cassette for rearranged κ light chain genes.

FIGURE 33
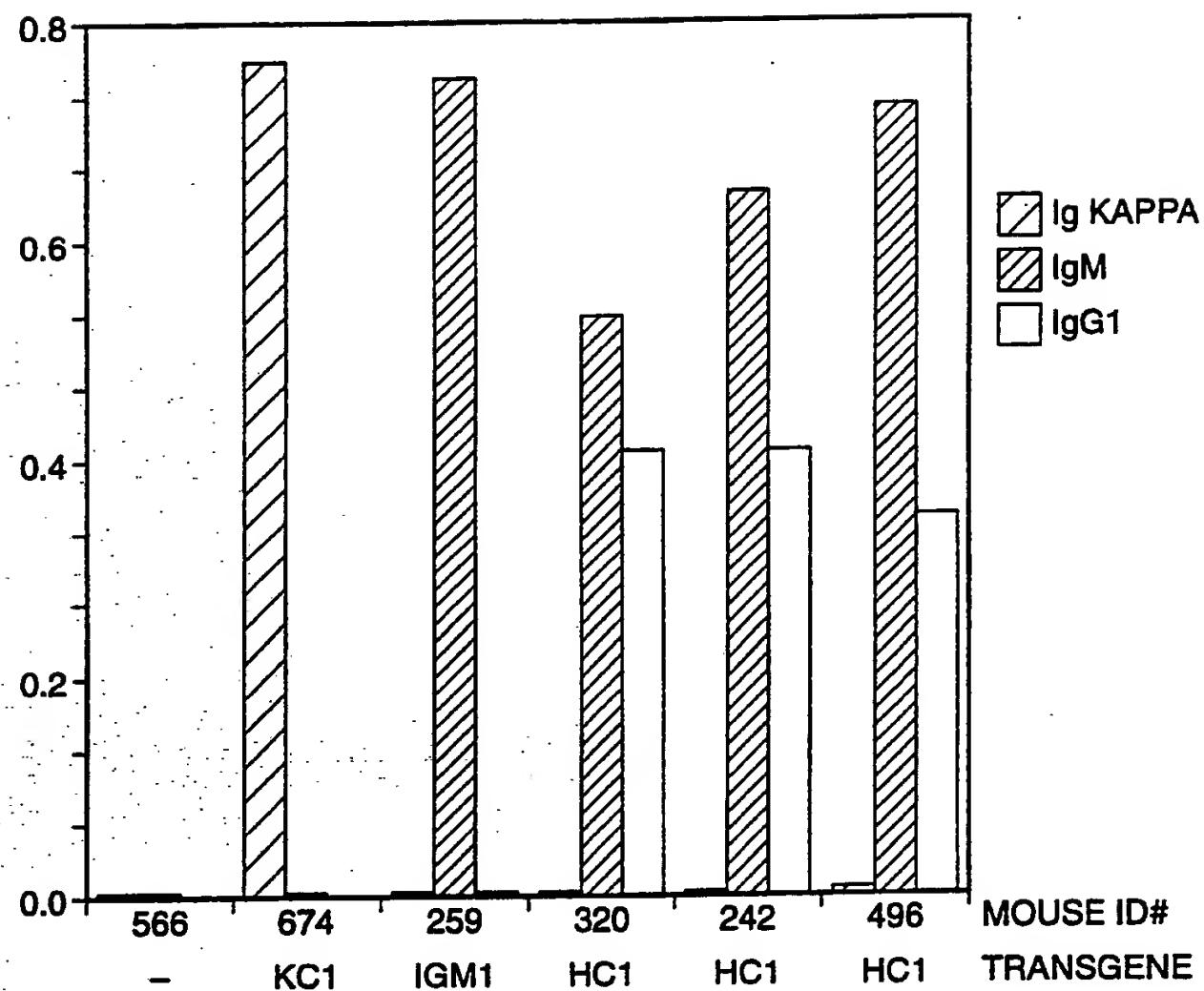
FIG. 34

37 / 47
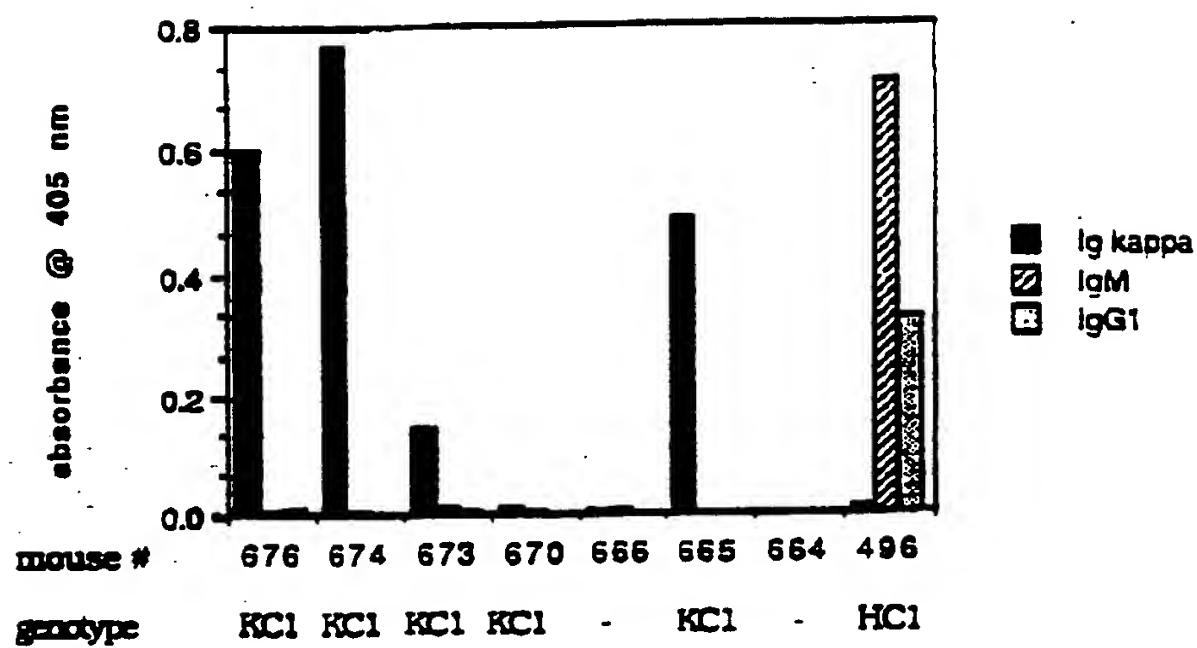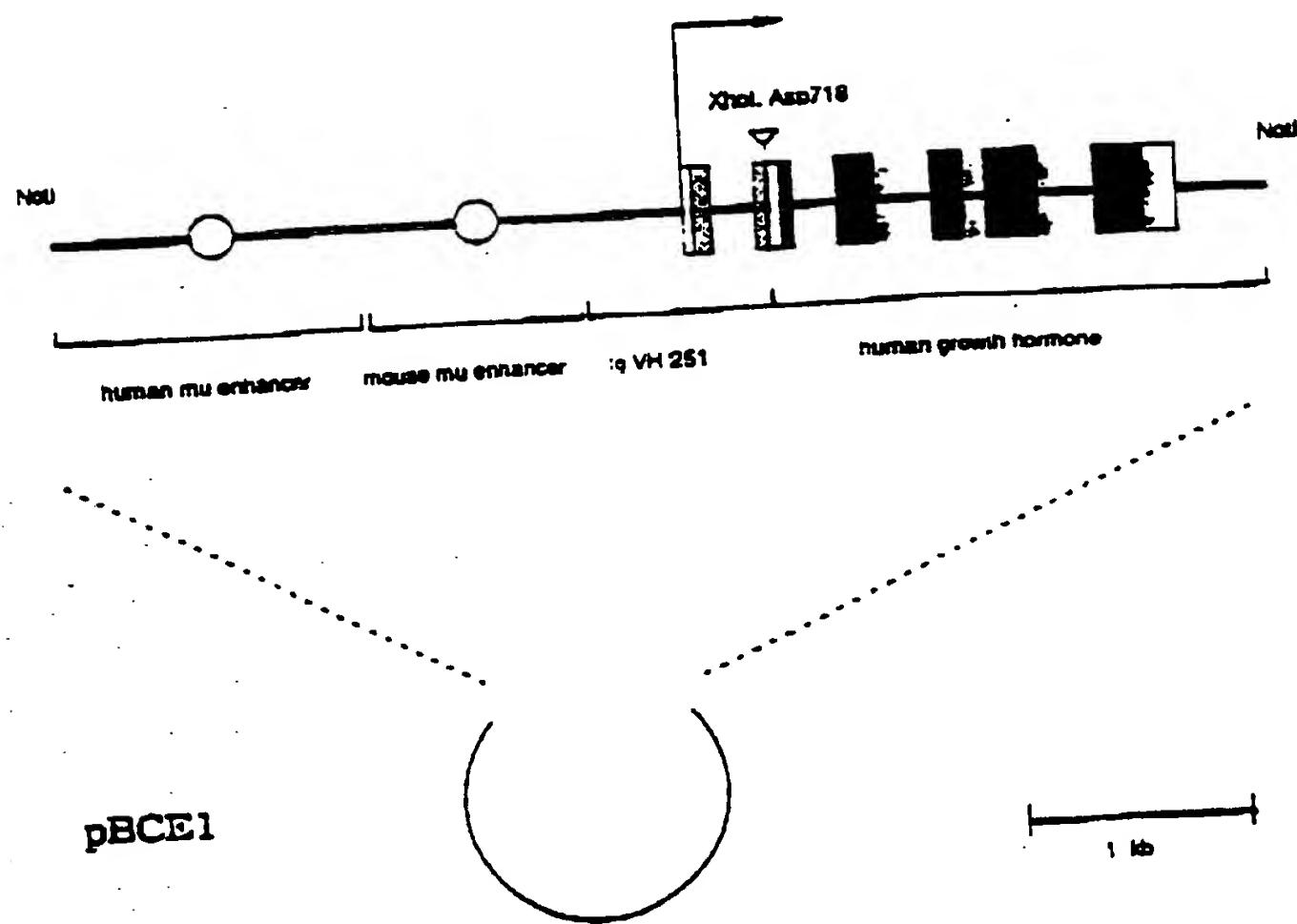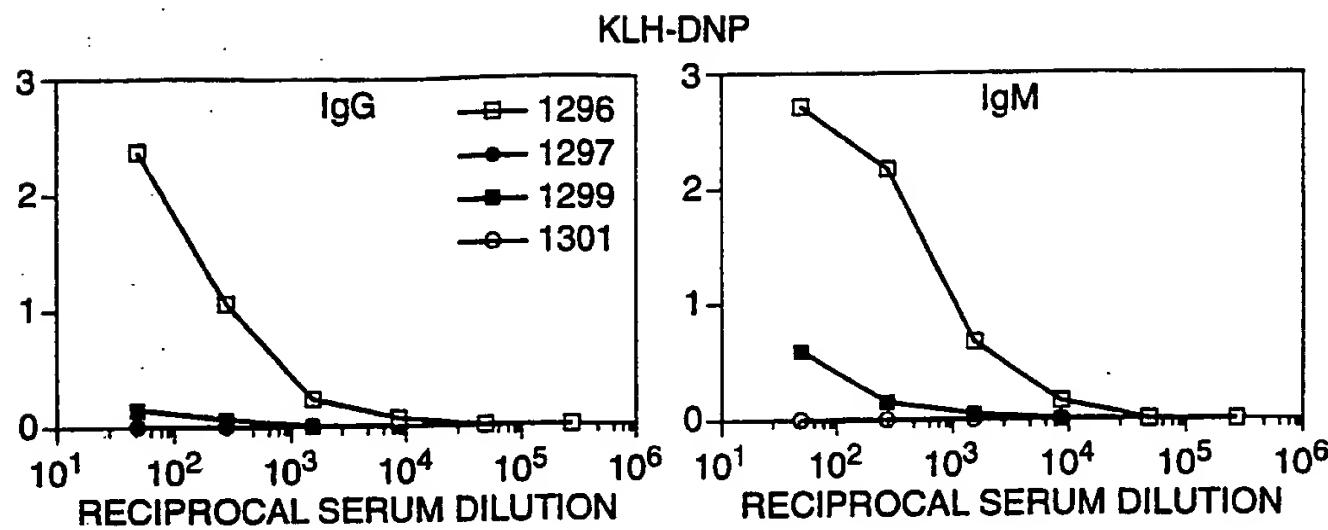
FIGURE 35
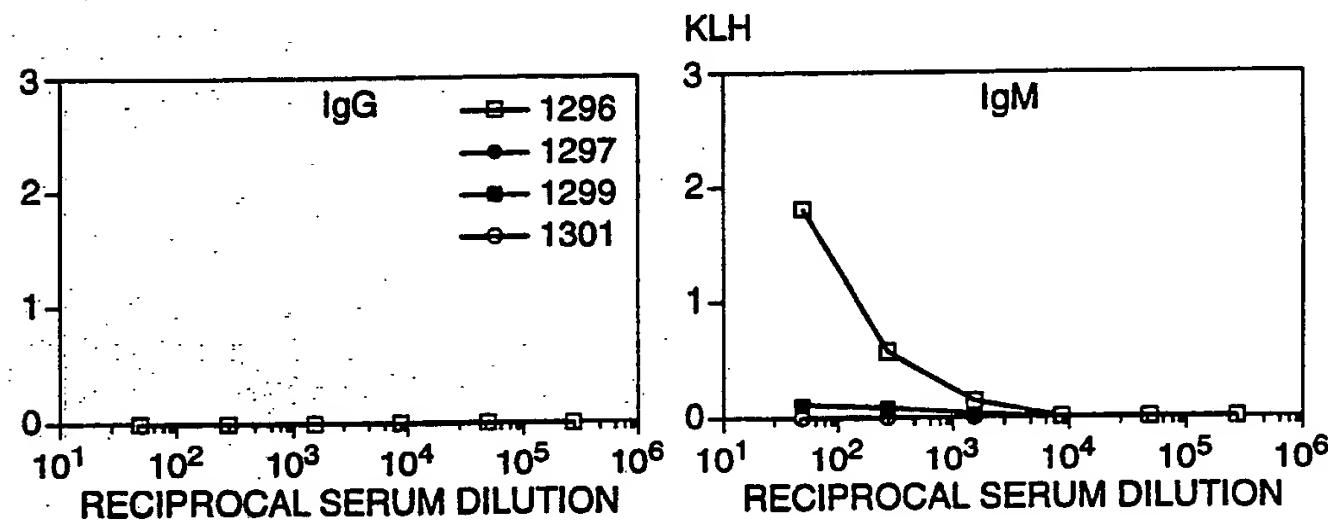
FIGURE 36

FIG. 37a

FIG. 37b

FIG. 37c

FIGURE 38
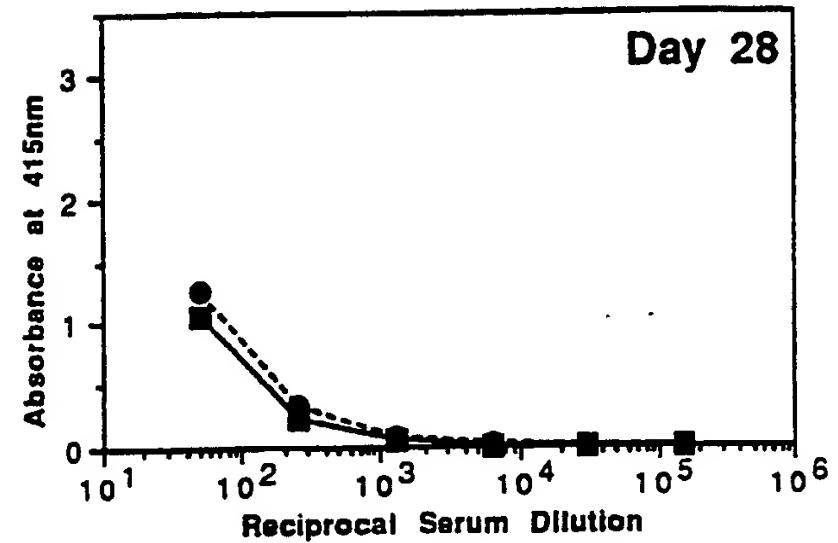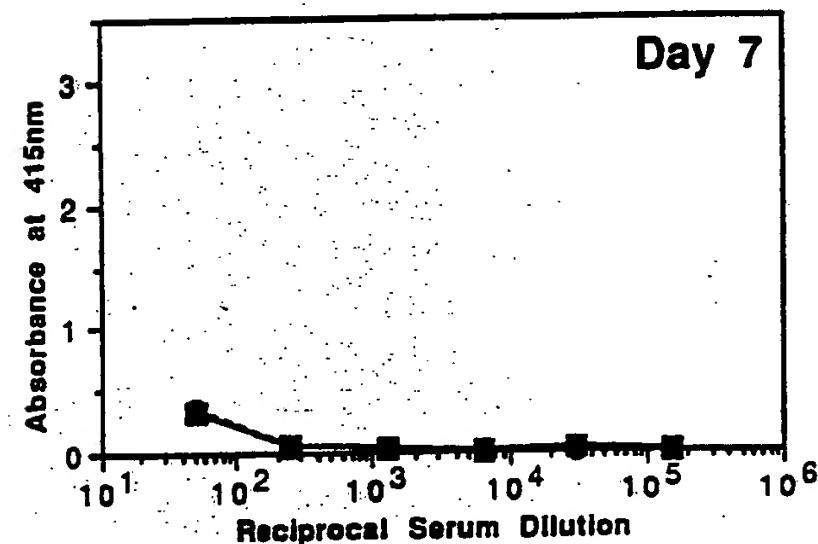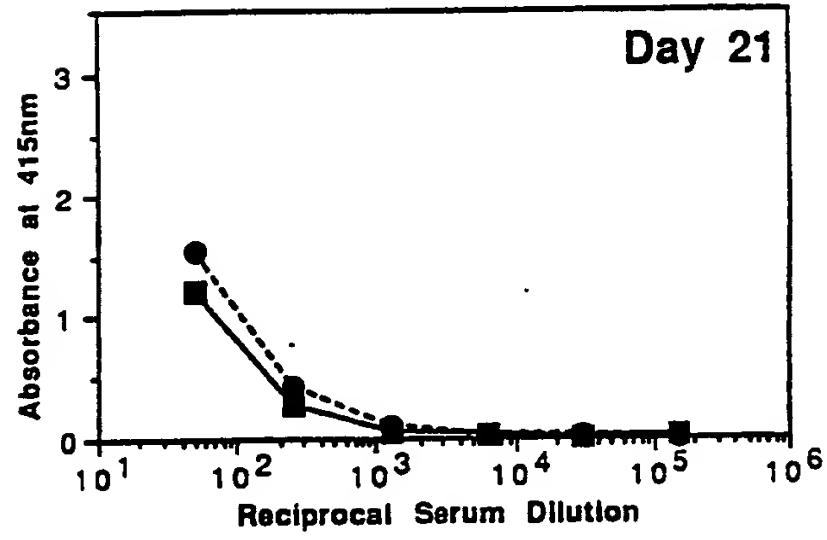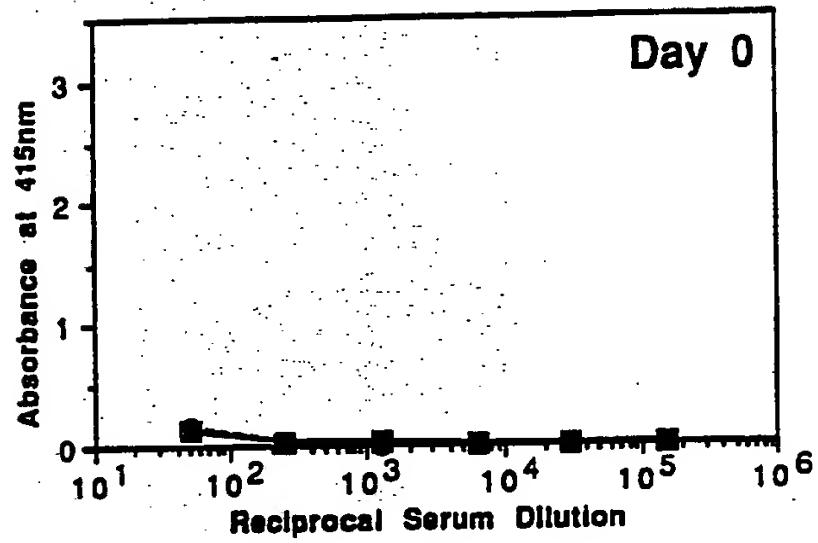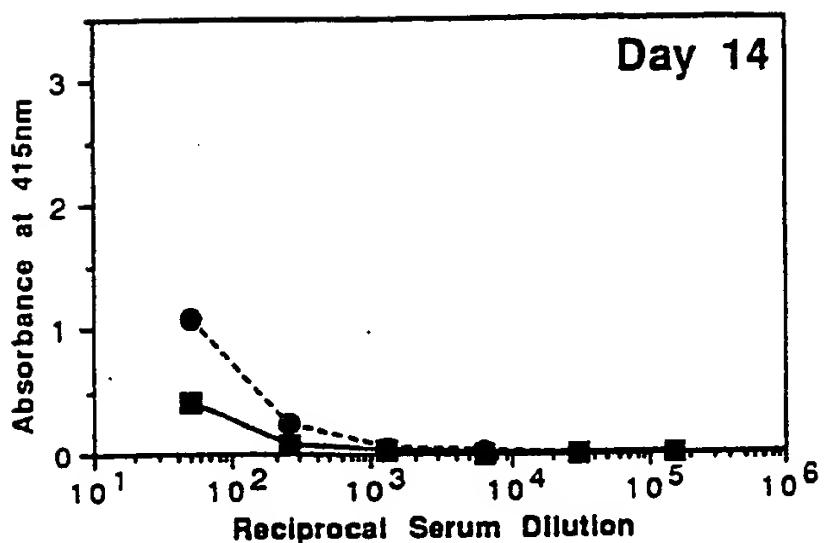
Human Mu heavy chain responses to human CEA determined by ELISA

FIGURE 39
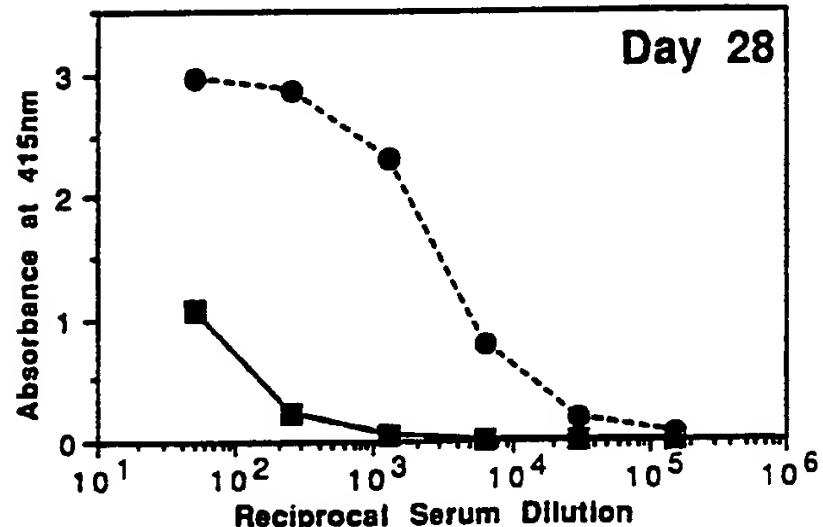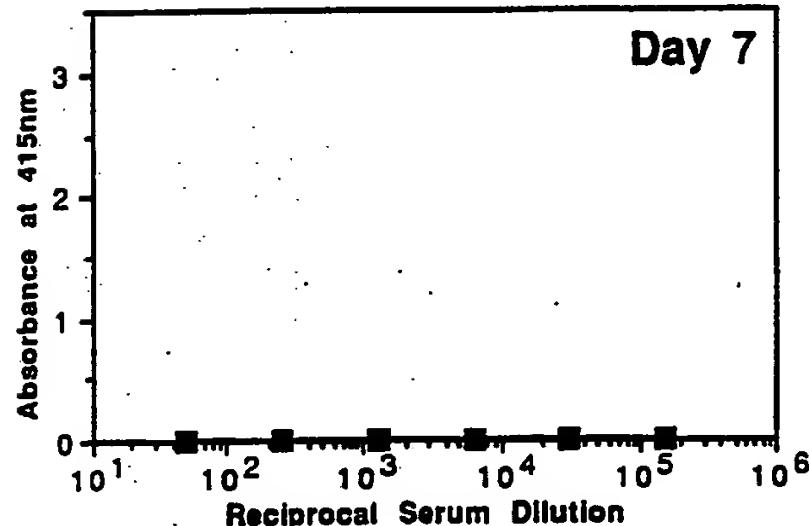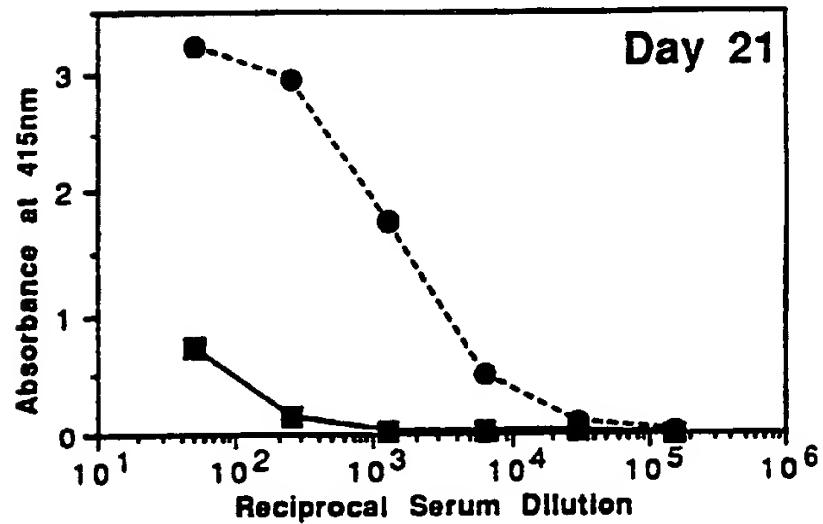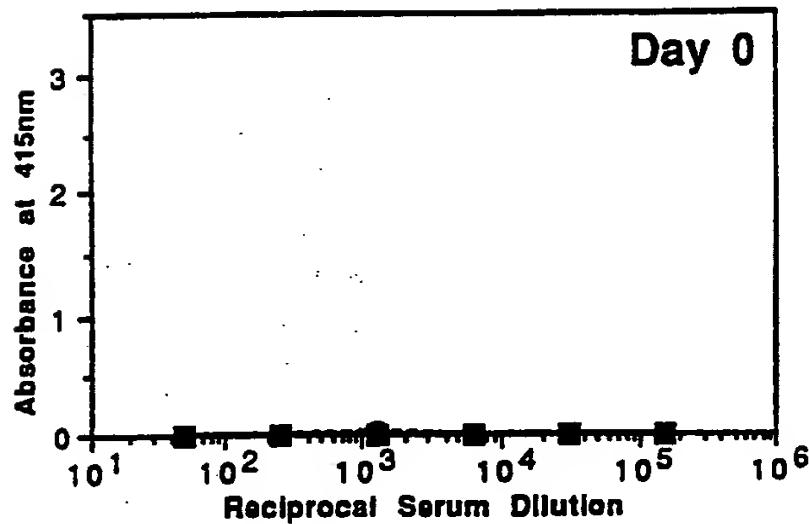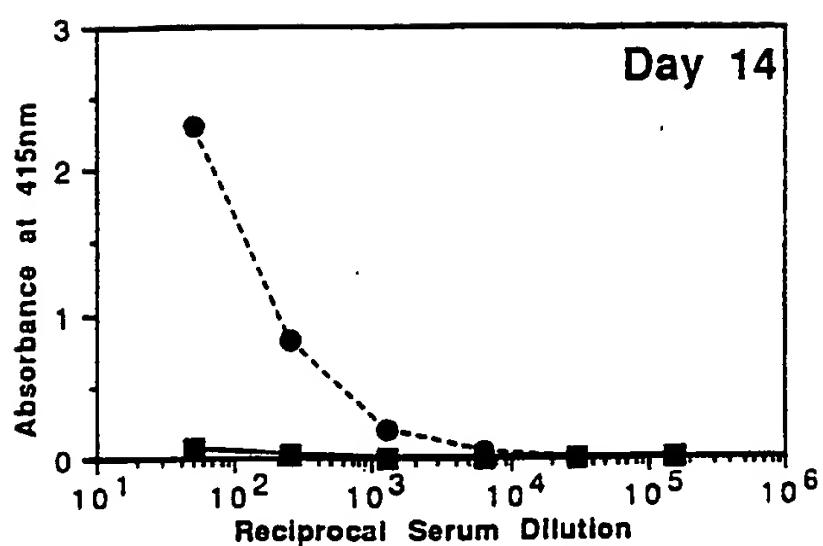
Human Gamma heavy chain responses to human CEA determined by ELISA

```
                                20                          30
G.L.   TCTCTGAAGATCTCCTGTAAGGGTTCTGGATACAGCTTTACCAGCTACTGGAT
         S  L  K  X  S  C  K  G  S  G  Y  S  F  T  S  V  N  X

32.

5.    ...........................C.........................

33.

10.   ....................................................

24.   ....................................................

34.

32.   ..............................................C...T.......
                                                   P   X

1.    ..............................................T...T.GA......
                                                       S  D

2.    ....................................................

3.    ....................................................

6.    ..............................................T...G.........
                                                        S

23.   ....................................................

30.   ....................................................

4.    .............................................C.A.....
                                                     Q

11.   ...........................................A..........
                                                  N

17.   ....................................................

27.   ....................................................

19.   ...................G......................T.S...T.......
                         R                          S

34.   ...................T.................A......T..T....
                         V                  L

36.   ....................................................

35.

25.   ..........................................A......C.......
                                                N        T

35.   ....................................................

36.

18.   ..........................................A....T..C.........
                                                N        X  T

22.   ..........................................C........T.......
                                                T

28.   ..........................................G...A.........
                                                A     N

33.   ....................................................
                                                └──────────┘
                                                    CDR I
```

FIG. 40a

```
                            40                        50
CGGCTGGGTGCGCCAGATGCCCGGGAAAGGCCTGGAGTGGATGGGGATCATCTATCCTGGT
  G   W   V   R   Q   N   P   G   K   G   L   E   W   N   G   X   X   V   P   G
```

FIG. 40b

CDR I   CDR II

```
                              60                                         70
         GACTCTGATACCAGATACAGCCCGTCCTTCCAAGGCCAGGTC            ACCATCTCAGC
          D  S  D  T  R  Y  S  P  S  F  Q  G  Q  V             T  X  S  A


         ...........................................G.........            ...........
                                                    R


         .........................................................            ...........


         .........................................................            ...........


         ........C......................G.........            T.......T
                                                                S       V


         .........................................................            ...........

         .........................................................            ...........

         .........................................................            ...........

         .........................................................            ...........

         .........................................................            ...........

         .........................................................            ...........

         .........................................................            ...........

         .........................................................            ...........

         .........................................................            ...........

         .......A.........F.....................            ...........
              E

         .........................................................            ...........

         .........................................................            ...........


         .........................................................            ...........

         .........................................................            ...........


         .........................................................            ...........


         .........................................................            ...........

         ......................C.................            ...........
                               P

         .........................................................            ...........
```

CDR II   FIG. 40c

```
                          80                              90
CGACAAGTCCATCAGCACCGCCTACCTGCAGTGGAGCAGCCTGAAGGCCTCGGACACCGCC
  D  K  S  X  S  T  A  Y  L  Q  W  S  S  L  K  A  S  D  T  A


........................................CG.........................
                                           E


..................................................................
...................................................G........GT.T.
                                                    R         S  V


..................................................................

............A.......T.........A..C...G....T...........
            N       F         N  T   E
..................................................................

..................................................................

............T.....................G.......................
                                  R
..................................................................

............T.....................................................

..................................................................

..................................................................

..................................................................

..................................................................

..................................................................

..........G.......................................................
           V

..................................................................


......C...........................................................
     N

..................................................................

..........A.......................................................
          N
```

FIG. 40d

```
ATGTATTACTGTGCGAGA
 N  Y  Y  C  A  R
                                                          TACTGGTAC
                                                           Y  W  Y
               .....................CAGGGGGGGGATA         ......
                                      Q  G  G  D           R
                                                             GCT
                                                             A
               .....................CATTGGctaaAtggggaT     ...
                                      N  D   N     D
               .....................CGGGattacgatattttgactggttattatGCG  ...
                                      R  D                        A
                                                             TAC
                                                             Y
               .....................GtggttcggggaTttattatT   ...
                                      V        F
               .....................GGgtattaTtatgAttcggggaCttattataagtctacCC  .
                                      G      D      T
               .....................ctaactggCCT            .
                                      G  L
               .....................CATCTT                 ...
                                      N  L
               .....................CATCTT                 ...
                                      N  L
               ..T..................CATCTT                 ...
                 X                    N  L
               ......CG.............CATCTT                 ...
                     R                N  L
               .....................CAAGG
                                      Q  G
               .....................CAAACT
                                      Q  T
               .....................CATggtatagcagcagctggtacGTGGTTCCGACCCCA
                                      N                    W  F  R  P  L
               .....................GCCggtataCcagcagctggtT
                                      A       T           F
               .....................CAGGGC                 ...
                                      Q  G
               .....................CAAAGGGG               ...
                                      Q  R  G
               .....................GGGATCGTGG             ...
                                      G  S  W
                                                          AACTGG
                                                           N  W
               .....................CTCCCCAATGACAGT        ...
                                      L  P  N  D  S
               .....................CGGGGGtactatggttcggggagttattat  ......
                                      R  G
                                                          TACTACTACTACTACGGT
                                                           Y  Y  Y  Y  Y  G
               .....................CATGagcagctggtacAGGGT  ...............
                                      N  E        Q  G
               .....................GATATGGGGGGGGCCTC      ......T.....T....
                                      D  N  G  G  A  S          F     F
               .....................                      C..C..G..........
                                                          H  H  D
               .....................CG                    ..........
                                      R
```

FIG. 40e   CDR III

```
TTCGATCTCTGGGGCCGTGGCACCCTGGTCACTGTCTCCTCAG
 P   D   L   W   G   R   G   T   L   V   T   V   S   S
 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

TTTGATATCTGGGGCCAAGGGACAATGGTCACCGTCTCTTCAG
 F   D   X   W   G   Q   G   T   N   V   T   V   S   S
 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

TTTGATATCTGGGGCCAAGGGACAATGGTCACCGTCTCTTCAG
 F   D   Y   N   G   Q   G   T   L   V   T   V   S   S
 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

 . . . . . . T . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

 . . . . A . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
         N

 . . . . . . . . . . . . . . . . . . . . A . . . . . . . . . . . . . . . . . . . . .
                                         Q
 . . . . . . . . . . . . . . . . . . . . A . . . . . . . . . . . . . . . . . . . . .
                                         Q
 . . . . . . . . . . . . . . . . . . . . A . . . . . . . . . . . . . . . . . . . . .
                                         Q
 . . . . . . . . . . . . . . . . . . . . A . . . . . . . . . . . . . . . . . . . . .
                                         Q

 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

                     G   P   G   N   P   G   H   R   L   L   S
 . . . . . . T . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
             P

 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

TTCGACCCCTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCAG
 F   D   P   W   G   Q   G   T   L   V   T   V   S   S
 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

ATGGACGTCTGGGGGCAAGGGACCACGGTCACCGTCTCCTCAG
 N   D   V   W   G   Q   G   T   T   V   T   V   S   S
 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
```

CDR III

FIG. 40f

```
TTTTCTGGCC TGACAACCAG GGTGGCGCAG GATGCTCAGT GCAGAGAGGA          50

AGAAGCAGGT GGTCTCTGCA GCTGGAAGCT CAGCTCCCAC CCAGCTGCTT         100

TGCATGTCCC TCCCAGCTGC CCTACCTTCC AGAGCCC ATA TCAAT GCCTG        150

TGTCAGAGCC CTGGGGAGGA ACTGCTCAGT TAGGACCCAG AGGGAACCAT         200
                                                     Me
GGAAGCCCCA GCTCAGCTTC TCTTCCTCCT GCTACTCTGG CTCCCAGgtg         250
tGluAlaPro AlaGlnLeuL euPheLeuLe uLeuLeuTrp LeuPro
aggggggaacc atgaggtggt tttgcacatt agtgaaaact cttgccacct         300

ctgctcagca agaaatataa ttaaaattca aagtatatca acaattttgg         350

ctctactcaa agacagttgg tttgatcttg attacatgag tgcatttctg         400

tttattttcc aatttcagAT ACCACCGGAG AAATTGTGTT GACACAGTCT         450
           Asp ThrThrGlyG luIleValLe uThrGlnSer
CCAGCCACCC TGTCTTTGTC TCCAGGGGAA AGAGCCACCC TCTCCTGCAG         500
ProAlaThrL euSerLeuSe rProGlyGlu ArgAlaThrL euSerCysAr
GGCCAGTCAG AGTGTTAGCA GCTACTTAGC CTGGTACCAA CAGAAACCTG         550
gAlaSerGln SerValSerS erTyrLeuAl aTrpTyrGln GlnLysProG
GCCAGGCTCC CAGGCTCCTC ATCTATGATG CATCCAACAG GGCCACTGGC         600
lyGlnAlaPr oArgLeuLeu IleTyrAspA laSerAsnAr gAlaThrGly
ATCCCAGCCA GGTTCAGTGG CAGTGGGTCT GGGACAGACT TCACTCTCAC         650
IleProAlaA rgPheSerGl ySerGlySer GlyThrAspP heThrLeuTh
CATCAGCAGC CTAGAGCCTG AAGATTTTGC AGTTTATTAC TGTCAGCAGC         700
rIleSerSer LeuGluProG luAspPheAl aValTyrTyr CysGlnGlnA
GTAGCAACTG GCCTCC ACA GTGATTCCAC ATGAA ACAAA AACC CCAACA        750
rgSerAsnTr pPro
AGACCATCAG TGTTTACTAG ATTATTATAC CAGCTGCTTC CTTTACAGAC         800

AGCTAGTGGG GT                                                  812
```

FIGURE 41

```
AGGGCGGCGC AGATGCTCAG TGCAGAGAGA AGAAACAGGT GGTCTCTGCA        50

GCTGGAAGCT CAGCTCCCAC CCCAGCTGCT TTGCATGTCC CTCCCAGCTG       100

CCCTACCTTC CAGAGCCCAT ATCAATGCCT GGGTCAGAGC TCTGGGGAGG       150

AACTGCTCAG TTAGGACCCA GACGGAACCA TGGAAGCCCC AGCGCAGCTT       200
                                M etGluAlaPr oAlaGlnLeu
CTCTTCCTCC TGCTACTCTG GCTCACAGgt gaggggaata tgaggtgtct       250
LeuPheLeuL euLeuLeuTr pLeuThr
ttgcacatca gtgaaaactc ctgccacctc tgctcagcaa gaaatataat        300

taaaattcaa aatagatcaa caattttggc tctactcaaa gacagtgggt        350

ttgattttga ttacatgagt gcatttctgt tttatttcca atttcagATA       400
                                                     AspT
CCACCGGAGA AATTGTGTTG ACACAGTCTC CAGCCACCCT GTCTTTGTCT       450
hrThrGlyGl uIleValLeu ThrGlnSerP roAlaThrLe uSerLeuSer
CCAGGGGAAA GAGCCACCCT CTCCTGCAGG GCCAGTCAGG GTGTTAGCAG       500
ProGlyGluA rgAlaThrLe uSerCysArg AlaSerGlnG lyValSerSe
CTACTTAGCC TGGTACCAGC AGAAACCTGG CCAGGCTCCC AGGCTCCTCA       550
rTyrLeuAla TrpTyrGlnG lnLysProGl yGlnAlaPro ArgLeuLeuI
TCTATGATGC ATCCAACAGG GCCACTGGCA TCCCAGCCAG GTTCAGTGGC       600
leTyrAspAl aSerAsnArg AlaThrGlyI leProAlaAr gPheSerGly
AGTGGGCCTG GGACAGACTT CACTCTCACC ATCAGCAGCC TAGAGCCTGA       650
SerGlyProG lyThrAspPh eThrLeuThr IleSerSerL euGluProGl
AGATTTTGCA GTTTATTACT GTCAGCAGCG TAGCAACTGG CATCCTACAG       700
uAspPheAla ValTyrTyrC ysGlnGlnAr gSerAsnTrp His
TGATTCCACA TGAAACAAAA ACCCAACAA GACCATCAGT GTTTACTAGA        750

TTATTATACC AGCTGCTTCC TTTACAGACA GCTAGTGGGG TGGCCACTCA       800

GTGTTAGCAT CTCAGCTCTA TTTGGCCATT TTGGAGTTCA AGTTGTCAAG       850

TCCAAAATTA CTTATGTTAG TCCATTGCAT CATACCATTT CAGTGTGGCT       900
```

FIGURE 42

45/47

CGGCCCAGC TGCTTTGCAT GTCCCTCCCA GCCGCCCTGC AGTCCAGAGC   50

CCATATCAAT GCCTGGGTCA GAGCTCTGGA GAAGAGCTGC TCAGTTAGGA   100

ACCCCAGAGG GAACCATGGA AACCCCAGCG CAGCTTCTCT TCCTCCTGCT   150
            MetGl uThrProAla GlnLeuLeuP heLeuLeuLe

ACTCTGGCTC CCAGgtgagg ggaacatggg atggttttgc atgtcagtga   200
uLeuTrpLeu Pro
aaaccctctc aagtcctgtt acctggcaac tctgctcagt caatacaata   250

attaaagctc aatataaagc aataattctg gctcttctgg gaagacaatg   300

ggtttgattt agattacatg ggtgactttt ctgttttatt tccaatctca   350

gATACCACCG GAGAAATTGT GTTGACGCAG TCTCCAGGCA CCCTGTCTTT   400
AspThrThrG lyGluIleVa lLeuThrGln SerProGlyT hrLeuSerLe
GTCTCCAGGG GAAAGAGCCA CCCTCTCCTG CAGGGCCAGT CAGAGTGTTA   450
uSerProGly GluArgAlaT hrLeuSerCy sArgAlaSer GlnSerValS
GCAGCAGCTA CTTAGCCTGG TACCAGCAGA AACCTGGCCA GGCTCCCAGG   500
erSerSerTy rLeuAlaTrp TyrGlnGlnL ysProGlyGl nAlaProArg
CTCCTCATCT ATGGTGCATC CAGCAGGGCC ACTGGCATCC CAGACAGGTT   550
LeuLeuIleT yrGlyAlaSe rSerArgAla ThrGlyIleP roAspArgPh
CAGTGGCAGT GGGTCTGGGA CAGACTTCAC TCTCACCATC AGCAGACTGG   600
eSerGlySer GlySerGlyT hrAspPheTh rLeuThrIle SerArgLeuG
AGCCTGAAGA TTTTGCAGTG TATTACTGTC AGCAGTATGG TAGCTCACCT   650
luProGluAs pPheAlaVal TyrTyrCysG lnGlnTyrGl ySerSerPro
CCCACAGTCA TTCAGCTTGA AACAAAAACC TCTGCAAGAC CTTCATTGTT   700

TACTAGATTA TACCAGCTGC TTCCTTTACA GATAGCTGCT GCAATGACAA   750

CTCAATTTAG CATCTCTCTC TGCTTGGGCA TTTTGGGGAT CTTAAAAAAG   800

TAATCCCTTG ATATATTTTT GACTCTGATT CCTGCATTTT TCCTCAGACC   850

AAGATGGACA GCCAGGTTTA AGCACAGTTT CACAGTAATG GCCACTGGAT   900

FIGURE 43

46/47

```
AAACACATTC TCTGCAGACA AATTTGAGCT ACCTTGATCT TACCTGGACA        50

GGTGGGGACA CTGAGCTGGT GCTGAGTTAC TCAGATGCGC CAGCTCTGCA       100

GCTGTGCCCA GCCTGCCCCA TCCCCTGCTC ATTTGCATGT TCCCAGAGCA       150

CAACCTCCTG CCCTGAAGCC TTATTAATAG GCTGGTCAGA CTTTGTGCAG       200

GAATCAGACC CAGTCAGGAC ACAGCATGGA CATGAGGGTC CTCGCTCAGC       250
                      MetAs pMetArgVal LeuAlaGlnL
TCCTGGGGCT CCTGCTGCTC TGTTTCCCAG gtaaggatgg agaacactag       300
euLeuGlyLe uLeuLeuLeu CysPhePro
cagtttactc agcccagggt gctcagtact gctttactat tcagggaaat       350

tctcttacaa catgattaat tgtgtggaca tttgttttta tgtttccaat       400

ctcagGTGCC AGATGTGACA TCCAGATGAC CCAGTCTCCA TCCTCACTGT       450
        GlyAla ArgCysAspI leGlnMetTh rGlnSerPro SerSerLeuS
CTGCATCTGT AGGAGACAGA GTCACCATCA CTTGTCGGGC GAGTCAGGGT       500
erAlaSerVa lGlyAspArg ValThrIleT hrCysArgAl aSerGlnGly
ATTAGCAGCT GGTTAGCCTG GTATCAGCAG AAACCAGAGA AAGCCCCTAA       550
IleSerSerT rpLeuAlaTr pTyrGlnGln LysProGluL ysAlaProLy
GTCCCTGATC TATGCTGCAT CCAGTTTGCA AAGTGGGGTC CCATCAAGGT       600
sSerLeuIle TyrAlaAlaS erSerLeuGl nSerGlyVal ProSerArgP
TCAGCGGCAG TGGATCTGGG ACAGATTTCA CTCTCACCAT CAGCAGCCTG       650
heSerGlySe rGlySerGly ThrAspPheT hrLeuThrIl eSerSerLeu
CAGCCTGAAG ATTTTGCAAC TTATTACTGC CAACAGTATA ATAGTTACCC       700
GlnProGluA spPheAlaTh rTyrTyrCys GlnGlnTyrA snSerTyrPr
ACCCACAGTC TTACACACCC AAACATAAAC CCCAGGGAA GCAGATGTGT       750
o
GAGGCTGGGC TGCCCAGCT GCTTCTCCTG ATGCCTCCAT CAGCTGAGAG       800

TGTTCCTCAG ATGCAGCCAC ACTCTGATGG TGTTGGTAGA TGGGGAC        847
```

FIGURE 44
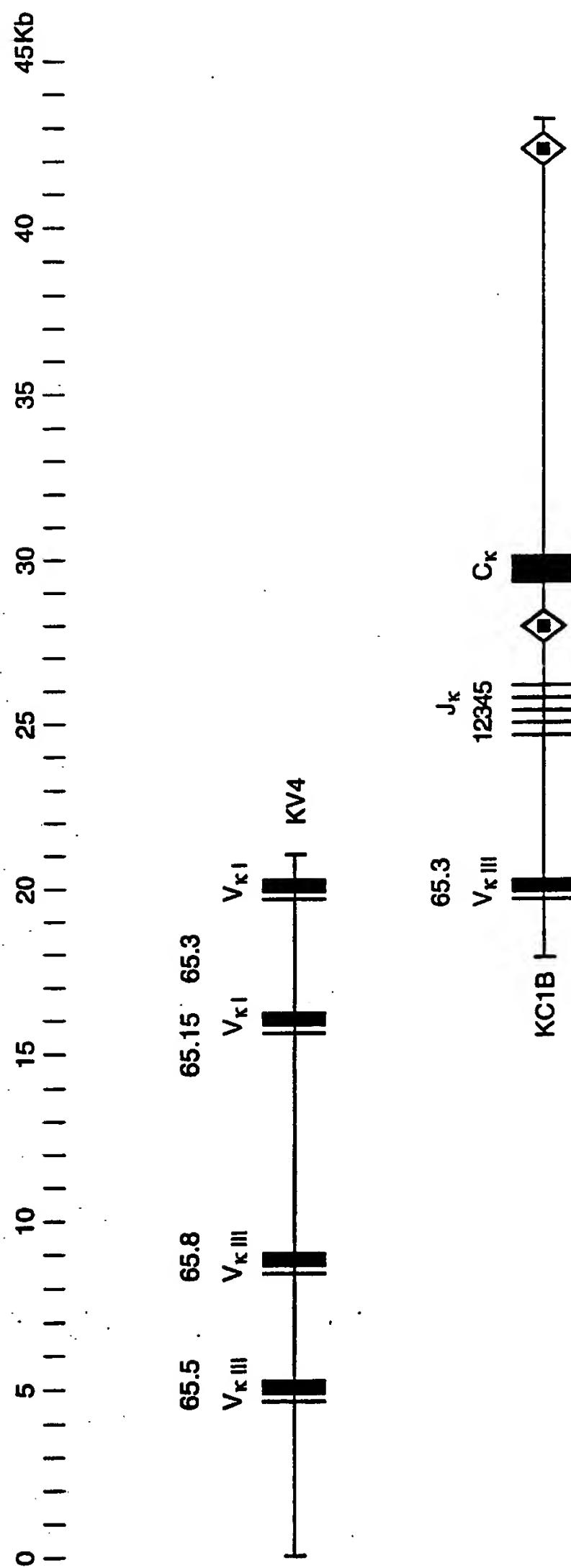
FIG. 45

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :C12N 15/00
US CL :800/2

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 800/2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS; heavy chain transgene
light chain transgene

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO,A, 90/04036 (Bruggemann et al) 19 April 1990, see entire article.	1-11, 12-26, 27, 28, 29-32, 33
Y	Nature, Volume 342, issued 23 November 1989, Zijlstra et al., "Germline transmission of a disrupted B2-microglobulin gene produced by homologous recombination in embryonic stem cells", pages 435-438, see entire article.	35-39, 43-53
Y	Nature, Volume 350, issued 04 April 1991, Kitamura et al., "A B cell-deficient mouse by targeted disruption of the membrane exon of the immunoglobulin u chain gene", pages 423-426, see entire article.	35-39, 54-58, 61-65

[X] Further documents are listed in the continuation of Box C.   [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance		
"E"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)		
		"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
03 MARCH 1993	07 APR 1993

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	SUZANNE ZISKA
Facsimile No. NOT APPLICABLE	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US92/10983

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
   because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
   because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
   because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest.
                        ☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet(1))(July 1992)∗

INTERNATIONAL SEARCH REPORT

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Chemical Reviews, Volume 90, No. 4, issued June 1990, Uhlmann et al., "Antisense Oligonucleotides: A new Therapeutic principle", pages 544-584, see entire article.	34, 40-42, 59, 60